

## University of Groningen

### Unveiling galaxy bias via the halo model, KiDS, and GAMA

Dvornik, Andrej; Hoekstra, Henk; Kuijken, Konrad; Schneider, Peter; Amon, Alexandra; Nakajima, Reiko; Viola, Massimo; Choi, Ami; Erben, Thomas; Farrow, Daniel J.

*Published in:*  
Monthly Notices of the Royal Astronomical Society

*DOI:*  
[10.1093/mnras/sty1502](https://doi.org/10.1093/mnras/sty1502)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2018

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Dvornik, A., Hoekstra, H., Kuijken, K., Schneider, P., Amon, A., Nakajima, R., Viola, M., Choi, A., Erben, T., Farrow, D. J., Heymans, C., Hildebrandt, H., Sifón, C., & Wang, L. (2018). Unveiling galaxy bias via the halo model, KiDS, and GAMA. *Monthly Notices of the Royal Astronomical Society*, 479(1), 1240-1259. <https://doi.org/10.1093/mnras/sty1502>

#### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

#### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Unveiling galaxy bias via the halo model, KiDS, and GAMA

Andrej Dvornik,<sup>1</sup>★ Henk Hoekstra,<sup>1</sup> Konrad Kuijken,<sup>1</sup> Peter Schneider,<sup>2</sup>  
Alexandra Amon,<sup>3</sup> Reiko Nakajima,<sup>2</sup> Massimo Viola,<sup>1</sup> Ami Choi,<sup>4</sup> Thomas Erben,<sup>2</sup>  
Daniel J. Farrow,<sup>5</sup> Catherine Heymans,<sup>3</sup> Hendrik Hildebrandt,<sup>2</sup> Cristóbal Sifón,<sup>6</sup> and  
Lingyu Wang<sup>7,8</sup>

<sup>1</sup>*Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, the Netherlands*

<sup>2</sup>*Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany*

<sup>3</sup>*SUPA, Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK*

<sup>4</sup>*Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*

<sup>5</sup>*Max-Planck-Institut für extraterrestrische Physik, Postfach 1312 Giessenbachstrasse, D-85741 Garching, Germany*

<sup>6</sup>*Department of Astrophysical Sciences, Peyton Hall, Princeton University, Princeton, NJ 08544, USA*

<sup>7</sup>*SRON Netherlands Institute for Space Research, Landleven 12, 9747 AD Groningen, the Netherlands*

<sup>8</sup>*Kapteyn Astronomical Institute, University of Groningen, Postbus 800, 9700 AV Groningen, the Netherlands*

Accepted 2018 June 6. Received 2018 June 6; in original form 2018 February 2

## ABSTRACT

We measure the projected galaxy clustering and galaxy–galaxy lensing signals using the Galaxy And Mass Assembly (GAMA) survey and Kilo-Degree Survey (KiDS) to study galaxy bias. We use the concept of non-linear and stochastic galaxy biasing in the framework of halo occupation statistics to constrain the parameters of the halo occupation statistics and to unveil the origin of galaxy biasing. The bias function  $\Gamma_{\text{gm}}(r_p)$ , where  $r_p$  is the projected comoving separation, is evaluated using the analytical halo model from which the scale dependence of  $\Gamma_{\text{gm}}(r_p)$ , and the origin of the non-linearity and stochasticity in halo occupation models can be inferred. Our observations unveil the physical reason for the non-linearity and stochasticity, further explored using hydrodynamical simulations, with the stochasticity mostly originating from the non-Poissonian behaviour of satellite galaxies in the dark matter haloes and their spatial distribution, which does not follow the spatial distribution of dark matter in the halo. The observed non-linearity is mostly due to the presence of the central galaxies, as was noted from previous theoretical work on the same topic. We also see that overall, more massive galaxies reveal a stronger scale dependence, and out to a larger radius. Our results show that a wealth of information about galaxy bias is hidden in halo occupation models. These models should therefore be used to determine the influence of galaxy bias in cosmological studies.

**Key words:** gravitational lensing: weak – methods: statistical – surveys – galaxies: haloes – dark matter – large-scale structure of Universe.

## 1 INTRODUCTION

In the standard cold dark matter and cosmological constant-dominated ( $\Lambda$ CDM) cosmological framework, galaxies form and reside within dark matter haloes, which themselves form from the highest density peaks in the initial Gaussian random density field (e.g. Mo, van den Bosch & White 2010 and references therein). In this case one expects that the spatial distribution of galaxies traces the spatial distribution of the underlying dark matter. Galaxies are however biased tracers of the underlying dark matter distribution,

because of the complexity of their evolution and formation (Davis et al. 1985; Dekel & Rees 1987; Cacciato et al. 2012). The relation between the distribution of galaxies and the underlying dark matter distribution, usually referred as *galaxy bias*, is thus important to understand in order to properly comprehend galaxy formation and interpret studies that use galaxies as tracers of the underlying dark matter, particularly for those trying to constrain cosmological parameters.

If such a relation can be described with a single number  $b$ , the galaxy bias is linear and deterministic. As galaxy formation is a complex process, it would be naive to assume that the relation between the dark matter density field and galaxies is a simple one, described only with a single number. Such a relation might be

★ E-mail: [dvornik@strw.leidenuniv.nl](mailto:dvornik@strw.leidenuniv.nl)

non-linear (the relation between a galaxy and matter density fields cannot be described with only a single number), scale dependent (the galaxy bias is different on the different scales studied), or stochastic (the biasing relation has an intrinsic scatter around the mean value). Numerous authors have presented various arguments for why simple linear and deterministic bias is highly questionable (Kaiser 1984; Davis et al. 1985; Dekel & Lahav 1999). Moreover, cosmological simulations and semi-analytical models suggest that galaxy bias takes a more complicated, non-trivial form (Wang et al. 2008; Zehavi et al. 2011).

Observationally, there have been many attempts to test if galaxy bias is linear and deterministic. There have been studies relying on clustering properties of different samples of galaxies (e.g. Wang et al. 2008; Zehavi et al. 2011), studies measuring high-order correlation statistics and ones directly comparing observed galaxy distribution fluctuations with the matter distribution fluctuations measured in numerical simulations (see Cacciato et al. 2012 and references therein). What is more, there have also been observations combining galaxy clustering with weak gravitational (galaxy–galaxy) lensing measurements (Hoekstra et al. 2002; Simon et al. 2007; Jullo et al. 2012; Buddendiek et al. 2016). The majority of the above observations have confirmed that galaxy bias is neither linear nor deterministic (Cacciato et al. 2012).

Even though the observational results are in broad agreement with theoretical predictions, until recently there was no direct connection between measurements and model predictions, mostly because the standard formalism used to define and predict the non-linearity and stochasticity of galaxy bias is hard to interpret in the framework of galaxy formation models. Cacciato et al. (2012) introduced a new approach that allows for intuitive interpretation of galaxy bias, that is directly linked to galaxy formation theory and various concepts therein. They reformulated the galaxy bias description (and the non-linearity and stochasticity of the relation between the galaxies and underlying dark matter distribution) presented by Dekel & Lahav (1999) using the formalism of halo occupation statistics. As galaxies are thought to live in dark matter haloes, halo occupation distributions (a prescription on how galaxies populate dark matter haloes) are a natural way to describe the galaxy–dark matter connection, and consequently the nature of galaxy bias. Combining the halo occupation distributions with the halo model (Peacock & Smith 2000; Seljak 2000; Cooray & Sheth 2002; van den Bosch et al. 2013; Mead et al. 2015; Wibking et al. 2017) allows us to compare observations to predictions of those models, which has the potential to unveil the *hidden factors* – sources of deviations from the linear and deterministic biasing (Cacciato et al. 2012). Recently Simon & Hilbert (2018) also showed that the halo model contains important information about galaxy bias. In this paper, however, we demonstrate how the stochasticity of galaxy bias arises from two different sources; the first is the relation between dark matter haloes and the underlying dark matter field, and the second is the manner in which galaxies populate dark matter haloes. As in Cacciato et al. (2012), we will focus on the second source of stochasticity, which indeed can be addressed using a halo model combined with halo occupation distributions.

The aim of this paper is to measure the galaxy bias using state of the art galaxy surveys and constrain the nature of it using the halo occupation distribution (HOD) formalism. The same formalism can provide us with insights on the sources of deviations from the linear and deterministic biasing and the results can be used in cosmological analyses using the combination of galaxy–galaxy lensing and galaxy clustering and those based on the cosmic shear measurements. In this paper, we make use of the predictions of Cacciato et al. (2012)

and apply them to the measurements provided by the imaging Kilo-Degree Survey (KiDS; Kuijken et al. 2015; de Jong et al. 2015), accompanied by the spectroscopic Galaxy And Mass Assembly (GAMA) survey (Driver et al. 2011) in order to get a grasp of the features of galaxy bias that can be measured using a combination of galaxy clustering and galaxy–galaxy lensing measurements with high precision.

The outline of this paper is as follows. In Section 2, we recap the galaxy biasing formulation of Cacciato et al. (2012). In Section 3, we introduce the halo model, its ingredients, and introduce the main observable, which is a combination of galaxy clustering and galaxy–galaxy lensing. In Section 4, we present the data and measurement methods used in our analysis. We present our galaxy biasing results in Section 5, together with comparison with simulations and discuss and conclude in Section 6. In the Appendix, we detail the calculation of the analytical covariance matrix, and provide full pairwise posterior distributions of our derived halo model parameters. We also provide a detailed derivation of the connection between the galaxy–matter correlation and the galaxy–galaxy lensing signal, explaining the use of two different definitions of the critical surface mass density in the literature. We highlight the key differences between our expressions and those found in several recent papers.

Throughout the paper we use the following cosmological parameters entering in the calculation of the distances and in the halo model (Planck Collaboration 2016):  $\Omega_m = 0.3089$ ,  $\Omega_\Lambda = 0.6911$ ,  $\sigma_8 = 0.8159$ ,  $n_s = 0.9667$ , and  $\Omega_b = 0.0486$ . We also use  $\bar{\rho}_m$  as the present day mean matter density of the Universe ( $\bar{\rho}_m = \Omega_{m,0} \rho_{\text{crit}}$ , where  $\rho_{\text{crit}} = 3H_0^2/(8\pi G)$ ) and the halo masses are defined as  $M = 4\pi r_\Delta^3 \bar{\rho}_m/3$  enclosed by the radius  $r_\Delta$  within which the mean density of the halo is  $\Delta$  times  $\bar{\rho}_m$ , with  $\Delta = 200$ ). All the measurements presented in the paper are in comoving units, and log and ln refer to the 10-based logarithm and the natural logarithm, respectively.

## 2 BIASING

This paper closely follows the biasing formalism presented in Cacciato et al. (2012), and we refer the reader to that paper for a thorough treatment of the topic. Here, we shortly recap the galaxy biasing formalism of Cacciato et al. (2012) and correct a couple of typos that we discovered during the study of his work. In this formalism the mean biasing function  $b(M)$  (the equivalent of the mean biasing function  $b(\delta_m)$  as defined by Dekel & Lahav 1999) is, using new variables: the number of galaxies in a dark matter halo,  $N$ , and the mass of a dark matter halo,  $M$ :

$$b(M) \equiv \frac{\bar{\rho}_m}{\bar{n}_g} \frac{\langle N|M \rangle}{M}, \quad (1)$$

where  $\bar{n}_g$  is the average number density of galaxies and  $N|M$  is the mean of the halo occupation distribution for a halo of mass  $M$ , defined as:

$$\langle N|M \rangle = \sum_{N=0}^{\infty} N P(N|M), \quad (2)$$

where  $P(N|M)$  is the halo occupation distribution. Note that in this case, the simple linear, deterministic biasing corresponds to:

$$N = \frac{\bar{n}_g}{\bar{\rho}_m} M, \quad (3)$$

which gives the expected value of  $b(M) = 1$ . As  $N$  is an integer and the quantities  $\bar{\rho}_m$ ,  $\bar{n}_g$ , and  $M$  are in general non-integer, it is clear

that in this formulation the linear, deterministic bias is unphysical. We define the moments of the bias function  $b(M)$  as

$$\hat{b} \equiv \frac{\langle b(M)M^2 \rangle}{\langle M^2 \rangle}, \quad (4)$$

and

$$\tilde{b}^2 \equiv \frac{\langle b^2(M)M^2 \rangle}{\langle M^2 \rangle}. \quad (5)$$

where  $\dots^1$  indicates an effective average (an integral over dark matter haloes) defined in the following form:

$$\langle x \rangle \equiv \int_0^\infty x n(M) dM, \quad (6)$$

where  $n(M)$  is the halo mass function and  $x$  is a property of the halo or galaxy population. In the case of linear bias,  $b(M)$  is a constant and hence  $\tilde{b}/\hat{b} = 1$ . The same ratio,  $\tilde{b}/\hat{b}$ , is the relevant measure of the non-linearity of the biasing relation (Dekel & Lahav 1999). Its deviation from unity is a sign of a non-linear galaxy bias. From equation (1) we can see that linear bias corresponds to halo occupation statistics for which  $N|M \propto M$ .

In the same manner, Cacciato et al. (2012) also define the random halo bias of a single halo of mass  $M$ , that contains  $N$  galaxies, as

$$\epsilon_N \equiv N - \langle N|M \rangle, \quad (7)$$

which, by definition, will have a zero mean when averaged over all dark matter haloes, i.e.  $\langle \epsilon_N | M \rangle = 0$ . This can be used to define the halo stochasticity function:

$$\sigma_b^2(M) \equiv \left( \frac{\bar{\rho}_m}{\bar{n}_g} \right)^2 \frac{\langle \epsilon_N^2 | M \rangle}{\langle M^2 \rangle}, \quad (8)$$

from which, after averaging over halo mass, one gets the stochasticity parameter:

$$\sigma_b^2 \equiv \left( \frac{\bar{\rho}_m}{\bar{n}_g} \right)^2 \frac{\langle \epsilon_N^2 \rangle}{\langle M^2 \rangle}. \quad (9)$$

If the stochasticity parameter  $\sigma_b = 0$ , then the galaxy bias is deterministic. In addition to the two bias moments  $\hat{b}$  and  $\tilde{b}$ , one can also define some other bias parameters, particularly the ratio of the variances  $b_{\text{var}}^2 \equiv \langle \delta_g^2 \rangle / \langle \delta_m^2 \rangle$  (Dekel & Lahav 1999; Cacciato et al. 2012). Using this definition and an HOD-based formulation, Cacciato et al. (2012) show that

$$b_{\text{var}}^2 = \left( \frac{\bar{\rho}_m}{\bar{n}_g} \right)^2 \frac{\langle N^2 \rangle}{\langle M^2 \rangle}, \quad (10)$$

where the averages are again calculated according to equation (6). As the bias parameter is sensitive to both non-linearity and stochasticity, the total variance of the bias  $b_{\text{var}}^2$  can also be written as

$$b_{\text{var}}^2 = \tilde{b}^2 + \sigma_b^2. \quad (11)$$

Combining equation (10) and (11), we find a relation for  $N^2$

$$\langle N^2 \rangle = \left( \frac{\bar{n}_g}{\bar{\rho}_m} \right)^2 [\tilde{b}^2 + \sigma_b^2] \langle M^2 \rangle. \quad (12)$$

We can compare this to the covariance, which is obtained directly from equations (1) and (3):

$$\langle NM \rangle = \frac{\bar{n}_g}{\bar{\rho}_m} \hat{b} \langle M^2 \rangle. \quad (13)$$

<sup>1</sup>Cacciato et al. (2012) used  $\sigma_M^2 \equiv \langle M^2 \rangle$  throughout the paper, and we decided to drop the  $\sigma_M^2$  for cleaner and more consistent equations.

From all the equations above, it also directly follows that one can define a linear correlation coefficient as:  $r \equiv \langle NM \rangle / [\langle N^2 \rangle \langle M^2 \rangle]$ , such that, combining equations (12) and (13),  $\hat{b}$  can be written as:  $\hat{b} = b_{\text{var}} r$ .

This enables us to consider some special cases. The discrete nature of galaxies does not allow us to have galaxy bias that is both linear and deterministic (Cacciato et al. 2012). Despite that, halo occupation statistics do allow bias that is linear and stochastic where

$$\hat{b} = \tilde{b} = b(M) = 1 \quad b_{\text{var}} = (1 + \sigma_b^2)^{1/2} \\ \sigma_b \neq 0 \quad r = (1 + \sigma_b^2)^{-1/2}, \quad (14)$$

or non-linear and deterministic

$$\hat{b} \neq \tilde{b} \neq 1 \quad b_{\text{var}} = \tilde{b} \\ \sigma_b = 0 \quad r = \hat{b}/\tilde{b} \neq 1. \quad (15)$$

### 3 HALO MODEL

To express the HOD, we use the halo model, a successful analytic framework used to describe the clustering of dark matter and its evolution in the Universe (Peacock & Smith 2000; Seljak 2000; Cooray & Sheth 2002; van den Bosch et al. 2013; Mead et al. 2015). The halo model provides an ideal framework to describe the statistical weak lensing signal around a selection of galaxies, their clustering, and cosmic shear signal. The halo model is built upon the statistical description of the properties of dark matter haloes (namely the average density profile, large scale bias, and abundance) as well as on the statistical description of the galaxies residing in them. The halo model allows us to unveil the *hidden* sources of bias stochasticity (Cacciato et al. 2012).

#### 3.1 Halo model ingredients

We assume that dark matter haloes are spherically symmetric, on average, and have density profiles,  $\rho(r|M) = M u_h(r|M)$ , that depend only on their mass  $M$ , and  $u_h(r|M)$  is the normalized density profile of a dark matter halo. Similarly, we assume that satellite galaxies in haloes of mass  $M$  follow a spherical number density distribution  $n_s(r|M) = N_s u_s(r|M)$ , where  $u_s(r|M)$  is the normalized density profile of satellite galaxies. Central galaxies always have  $r = 0$ . We assume that the density profile of dark matter haloes follows an NFW profile (Navarro, Frenk & White 1997). Since centrals and satellites are distributed differently, we write the galaxy–galaxy power spectrum as

$$P_{\text{gg}}(k) = f_c^2 P_{\text{cc}}(k) + 2f_c f_s P_{\text{cs}}(k) + f_s^2 P_{\text{ss}}(k), \quad (16)$$

while the galaxy–dark matter cross power spectrum is given by

$$P_{\text{gm}}(k) = f_c P_{\text{cm}}(k) + f_s P_{\text{sm}}(k). \quad (17)$$

Here,  $f_c = \bar{n}_c/\bar{n}_g$  and  $f_s = \bar{n}_s/\bar{n}_g = 1 - f_c$  are the central and satellite fractions, respectively, and the average number densities  $\bar{n}_g$ ,  $\bar{n}_c$ , and  $\bar{n}_s$  follow from:

$$\bar{n}_x = \int_0^\infty \langle N_x | M \rangle n(M) dM, \quad (18)$$

where ‘x’ stands for ‘g’ (for galaxies), ‘c’ (for centrals), or ‘s’ (for satellites) and  $n(M)$  is the halo mass function in the following form:

$$n(M) = \frac{\bar{\rho}_m}{M^2} \nu f(\nu) \frac{d \ln \nu}{d \ln M}, \quad (19)$$

with  $\nu = \delta_c/\sigma(M)$ , where  $\delta_c$  is the critical overdensity for spherical collapse at redshift  $z$ , and  $\sigma(M)$  is the mass variance. For  $f(\nu)$ , we use

the form presented in Tinker et al. (2010). In addition, it is common practice to split two-point statistics into a 1-halo term (both points are located in the same halo) and a 2-halo term (the two points are located in different haloes). The 1-halo terms are as follows:

$$P_{cc}^{1h}(k) = \frac{1}{\bar{n}_c}, \quad (20)$$

$$P_{ss}^{1h}(k) = \beta \int_0^\infty \mathcal{H}_s^2(k, M) n(M) dM, \quad (21)$$

and all other terms are given by

$$P_{xy}^{1h}(k) = \int_0^\infty \mathcal{H}_x(k, M) \mathcal{H}_y(k, M) n(M) dM. \quad (22)$$

Here ‘x’ and ‘y’ are either ‘c’ (for central), ‘s’ (for satellite), or ‘m’ (for matter),  $\beta$  is a Poisson parameter that arises from considering a scatter in the number of satellite galaxies at fixed halo mass [in this case a free parameter – we define the  $\beta$  in detail using equations (40)–(42)] and we have defined

$$\mathcal{H}_m(k, M) = \frac{M}{\bar{\rho}_m} \tilde{u}_h(k|M), \quad (23)$$

$$\mathcal{H}_c(k, M) = \frac{\langle N_c|M \rangle}{\bar{n}_c}, \quad (24)$$

and

$$\mathcal{H}_s(k, M) = \frac{\langle N_s|M \rangle}{\bar{n}_s} \tilde{u}_s(k|M), \quad (25)$$

with  $\tilde{u}_h(k|M)$  and  $\tilde{u}_s(k|M)$  the Fourier transforms of the halo density profile and the satellite number density profile, respectively, both normalized to unity [ $\tilde{u}(k=0|M)=1$ ]. The various 2-halo terms are given by

$$P_{xy}^{2h}(k) = P_{lin}(k) \int_0^\infty dM_1 \mathcal{H}_x(k, M_1) b_h(M_1) n(M_1) \times \int_0^\infty dM_2 \mathcal{H}_y(k, M_2) b_h(M_2) n(M_2), \quad (26)$$

where  $P_{lin}(k)$  is the linear power spectrum, obtained using the Eisenstein & Hu (1998) transfer function, and  $b_h(M, z)$  is the halo bias function. Note that in this formalism, the matter–matter power spectrum simply reads:

$$P_{mm}(k) = P_{mm}^{1h}(k) + P_{mm}^{2h}(k). \quad (27)$$

The two-point correlation functions corresponding to these power-spectra are obtained by simple Fourier transformation:

$$\xi_{xy}(r) = \frac{1}{2\pi^2} \int_0^\infty P_{xy}(k) \frac{\sin kr}{kr} k^2 dk, \quad (28)$$

For the halo bias function,  $b_h$ , we use the fitting function from Tinker et al. (2010), as it was obtained using the same numerical simulation from which the halo mass function was obtained. We have adopted the parametrization of the concentration–mass relation, given by Duffy et al. (2008):

$$c(M, z) = 10.14 A_c \left[ \frac{M}{(2 \times 10^{12} M_\odot/h)} \right]^{-0.081} (1+z)^{-1.01}, \quad (29)$$

with a free normalization  $A_c$  that accounts for the theoretical uncertainties in the concentration–mass relation due to discrepancies in the numerical simulations (mostly resolution and cosmologies) from which this scaling is usually inferred (Viola et al. 2015). We allow for additional normalization  $A_s$  for satellites, such that

$$c_s(M, z) = A_s c(M, z), \quad (30)$$

which governs how satellite galaxies are spatially distributed inside a dark matter halo and tests the assumption of satellite galaxies following the density distribution of the dark matter haloes. If  $A_s \neq 1$ , the galaxy bias will vary on small scales, as demonstrated by Cacciato et al. (2012).

### 3.2 Conditional stellar mass function

In order to constrain the cause for the stochasticity, non-linearity and scale dependence of galaxy bias, we model the halo occupation statistics using the Conditional Stellar Mass Function (CSMF, heavily motivated by Yang, Mo & van den Bosch 2008; Cacciato et al. 2009, 2013; Wang et al. 2013; van Uitert et al. 2016). The CSMF,  $\Phi(M_\star|M)$ , specifies the average number of galaxies of stellar mass  $M_\star$  that reside in a halo of mass  $M$ . In this formalism, the halo occupation statistics of central galaxies are defined via the function:

$$\Phi(M_\star|M) = \Phi_c(M_\star|M) + \Phi_s(M_\star|M). \quad (31)$$

In particular, the CSMF of central galaxies is modelled as a lognormal

$$\Phi_c(M_\star|M) = \frac{1}{\sqrt{2\pi} \ln(10) \sigma_c M_\star} \exp \left[ -\frac{\log(M_\star/M_c^*)^2}{2 \sigma_c^2} \right], \quad (32)$$

and the satellite term as a modified Schechter function

$$\Phi_s(M_\star|M) = \frac{\phi_s^*}{M_s^*} \left( \frac{M_\star}{M_s^*} \right)^{\alpha_s} \exp \left[ -\left( \frac{M_\star}{M_s^*} \right)^2 \right], \quad (33)$$

where  $\sigma_c$  is the scatter between stellar mass and halo mass and  $\alpha_s$  governs the power-law behaviour of satellite galaxies. Note that  $M_c^*$ ,  $\sigma_c$ ,  $\phi_s^*$ ,  $\alpha_s$ , and  $M_s^*$  are, in principle, all functions of halo mass  $M$ . We assume that  $\sigma_c$  and  $\alpha_s$  are independent of the halo mass  $M$ . Inspired by Yang et al. (2008), we parametrize  $M_c^*$ ,  $M_s^*$ , and  $\phi_s^*$  as

$$M_c^*(M) = M_0 \frac{(M/M_1)^{\gamma_1}}{[1 + (M/M_1)^{\gamma_1 - \gamma_2}]^{\gamma_1 - \gamma_2}}, \quad (34)$$

$$M_s^*(M) = 0.56 M_c^*(M), \quad (35)$$

and

$$\log[\phi_s^*(M)] = b_0 + b_1 (\log m_{12}), \quad (36)$$

where  $m_{12} = M/(10^{12} M_\odot/h)$ . The factor of 0.56 is also inspired by Yang et al. (2008) and further tests by van Uitert et al. (2016) showed that using this assumption does not significantly affect the results. We can see that the stellar to halo mass relation for  $M \ll M_1$  behaves as  $M_c^* \propto M^{\gamma_1}$  and for  $M \gg M_1$ ,  $M_c^* \propto M^{\gamma_2}$ , where  $M_1$  is a characteristic mass scale and  $M_0$  is a normalization. Here,  $\gamma_1$ ,  $\gamma_2$ ,  $b_0$ , and  $b_1$  are all free parameters.

From the CSMF, it is straightforward to compute the halo occupation numbers. For example, the average number of galaxies with stellar masses in the range  $M_{\star,1} \leq M_\star \leq M_{\star,2}$  is thus given by:

$$\langle N|M \rangle = \int_{M_{\star,1}}^{M_{\star,2}} \Phi(M_\star|M) dM_\star. \quad (37)$$

The distinction we have made here, by splitting galaxies into centrals or satellites, is required to illustrate the main source of non-linearity and scale dependence of galaxy bias (see results in Section 5). To explore this, we follow Cacciato et al. (2012), and define the random halo biases following similar procedure as in equation (7)

$$\varepsilon_c \equiv N_c - \langle N_c|M \rangle \quad \text{and} \quad \varepsilon_s \equiv N_s - \langle N_s|M \rangle, \quad (38)$$



and the halo stochasticity functions for centrals and satellites are given by

$$\begin{aligned}\langle \varepsilon_c^2 | M \rangle &= \sum_{N_c=0}^{\infty} (N_c - \langle N_c | M \rangle)^2 P(N_c | M) \\ &= \langle N_c^2 | M \rangle - \langle N_c | M \rangle^2 \\ &= \langle N_c | M \rangle - \langle N_c | M \rangle^2,\end{aligned}\quad (39)$$

$$\begin{aligned}\langle \varepsilon_s^2 | M \rangle &= \sum_{N_s=0}^{\infty} (N_s - \langle N_s | M \rangle)^2 P(N_s | M) \\ &= \langle N_s^2 | M \rangle - \langle N_s | M \rangle^2,\end{aligned}\quad (40)$$

where we have used the fact that  $\langle N_c^2 | M \rangle = \langle N_c | M \rangle$ , which follows from the fact that  $N_c$  is either zero or unity. We can see that central galaxies only contribute to the stochasticity if  $N_c | M < 1$ . If  $N_c | M = 1$ , then the HOD is deterministic and the stochasticity function  $\langle \varepsilon_c^2 | M \rangle = 0$ . The CSMF, however, only specifies the first moment of the halo occupation distribution  $P(N | M)$ . For central galaxies this is not a problem, as  $\langle N_c^2 | M \rangle = \langle N_c | M \rangle$ . For satellite galaxies, we use that

$$\langle N_s^2 | M \rangle = \beta(M) \langle N_s | M \rangle^2 + \langle N_s | M \rangle, \quad (41)$$

where  $\beta(M)$  is the mass-dependent Poisson parameter defined as

$$\beta(M) \equiv \frac{\langle N_s(N_s - 1) | M \rangle}{\langle N_s | M \rangle^2}, \quad (42)$$

which is unity if  $P(N_s | M)$  is given by a Poisson distribution, larger than unity if the distribution is wider than a Poisson distribution (also called super-Poissonian distribution) or smaller than unity if the distribution is narrower than a Poisson distribution (also called sub-Poissonian distribution). If  $\beta(M)$  is unity, then equation (40) takes a simple form  $\langle \varepsilon_s^2 | M \rangle = \langle N_s | M \rangle$ .

In what follows we limit ourselves to cases in which  $\beta(M)$  is independent of halo mass, i.e.  $\beta(M) = \beta$ , and we treat  $\beta$  as a free parameter.

Even without an application to the data, we can already learn a lot about the nature of galaxy bias from combining the HOD and halo model approaches to galaxy biasing as described in Section 2. As realistic HODs (as formulated above) differ strongly from the simple scaling  $N | M \propto M$  (equation 3), which gives the linear and deterministic galaxy bias), they will inherently predict a galaxy bias that is strongly non-linear. Moreover, this seems to be mostly the consequence of central galaxies for which  $N_c | M$  never follows a power law. Even the satellite occupation distribution  $N_s | M$  is never close to the power-law form, due to a cut-off at the low mass end, as galaxies at certain stellar mass require a minimum mass for their host halo (Cacciato et al. 2012 see also fig. 2 therein). Given the behaviour of the halo model and the HOD, the stochasticity of the galaxy bias could most strongly arise from the non-zero  $\sigma_c$  in equation (32) and the possible non-Poissonian nature of the satellite galaxy distribution for less massive galaxies. For more massive galaxies the main source of stochasticity can be shot noise, which dominates the stochasticity function,  $\sigma_b$  in equation (9), when the number density of galaxies is small. We use those free parameters of the HOD in a fit to the data (see Section 4), to constrain the cause for the stochasticity, non-linearity, and scale dependence of galaxy bias.

**Table 1.** Overview of the median stellar masses of galaxies, median redshifts, and number of galaxies/lenses in each selected bin, which are indicated in the second column. Stellar masses are given in units of  $[\log(M_*/[M_\odot/h^2])]$ .

Sample	Range	$M_{*,\text{med}}$	$z_{\text{med}}$	# of lenses
Bin 1	(10.3, 10.6]	10.46	0.244	26 224
Bin 2	(10.6, 10.9]	10.74	0.284	20 452
Bin 3	(10.9, 12.0]	11.13	0.318	10 178

### 3.3 Projected functions

We can project the 3D bias functions as defined by Dekel & Lahav (1999) and Cacciato et al. (2012) into 2D, projected analogues, which are more easily accessible observationally. We start by defining the matter–matter, galaxy–matter, and galaxy–galaxy projected surface densities as

$$\Sigma_{xy}(r_p) = 2\bar{\rho}_m \int_{r_p}^{\infty} \xi_{xy}(r) \frac{r dr}{\sqrt{r^2 - r_p^2}}, \quad (43)$$

where ‘x’ and ‘y’ stand either for ‘g’ or ‘m’, and  $r_p$  is the *projected* separation, with the change from standard line-of-sight integration to the integration along the projected separation using an Abel transformation. We also define  $\bar{\Sigma}_{xy}(< r_p)$  as its average inside  $r_p$ :

$$\bar{\Sigma}_{xy}(< r_p) = \frac{2}{r_p^2} \int_0^{r_p} \Sigma_{xy}(R') R' dR', \quad (44)$$

which we use to define the excess surface densities (ESD)

$$\Delta \Sigma_{xy}(r_p) = \bar{\Sigma}_{xy}(< r_p) - \Sigma_{xy}(r_p). \quad (45)$$

We include the contribution of the stellar mass of galaxies to the lensing signal as a point mass approximation, which we can write as

$$\Delta \Sigma_{\text{gm}}^{\text{pm}}(r_p) = \frac{M_{*,\text{med}}}{\pi r_p^2}, \quad (46)$$

where  $M_{*,\text{med}}$  is the median stellar mass of the selected galaxies obtained directly from the GAMA catalogue (Taylor et al. 2011; see Section 4.1 and Table 1 for more details). This stellar mass contribution is fixed by each of our samples. According to the checks performed, the inclusion of the stellar mass contribution to the lensing signal does not affect our conclusions.

The obtained projected surface densities can subsequently be used to define the projected, 2D analogues of the 3D bias functions ( $b_g^{\text{3D}}$ ,  $\mathcal{R}_{\text{gm}}^{\text{3D}}$ , and  $\Gamma_{\text{gm}}^{\text{3D}}$ , Dekel & Lahav 1999; Cacciato et al. 2012) as

$$b_g(r_p) \equiv \sqrt{\frac{\Delta \Sigma_{\text{gg}}(r_p)}{\Delta \Sigma_{\text{mm}}(r_p)}}, \quad (47)$$

$$\mathcal{R}_{\text{gm}}(r_p) \equiv \frac{\Delta \Sigma_{\text{gm}}(r_p)}{\sqrt{\Delta \Sigma_{\text{gg}}(r_p) \Delta \Sigma_{\text{mm}}(r_p)}}, \quad (48)$$

and

$$\Gamma_{\text{gm}}(r_p) \equiv \frac{b_g(r_p)}{\mathcal{R}_{\text{gm}}(r_p)} = \frac{\Delta \Sigma_{\text{gg}}(r_p)}{\Delta \Sigma_{\text{gm}}(r_p)}. \quad (49)$$

In what follows we shall refer to these as the ‘projected bias functions’.

In the case of the galaxy–dark matter cross correlation, the excess surface density  $\Delta \Sigma_{\text{gm}}(r_p) = \gamma_t(r_p) \Sigma_{\text{cr,com}}$ , where  $\gamma_t(r_p)$  is the tangential shear, which can be measured observationally using

galaxy–galaxy lensing, and  $\Sigma_{\text{cr,com}}$  is the comoving critical surface mass density:<sup>2</sup>

$$\Sigma_{\text{cr,com}} = \frac{c^2}{4\pi G(1+z_l)^2} \frac{D(z_s)}{D(z_l)D(z_l, z_s)}, \quad (50)$$

where  $D(z_l)$  is the angular diameter distance to the lens,  $D(z_l, z_s)$  is the angular diameter distance between the lens and the source, and  $D(z_s)$  is the angular diameter distance to the source. In Appendix B, we discuss the exact derivation of equation (50) and the implications of using different coordinates. In the case of the galaxy–galaxy autocorrelation we can write that

$$\Delta \Sigma_{\text{gg}}(r_p) = \bar{\rho}_m \left[ \frac{2}{r_p^2} \int_0^{r_p} w_p(R') R' dR' - w_p(r_p) \right], \quad (51)$$

where  $w_p(r_p)$  is the projected galaxy correlation function, and  $w_p(r_p) = \Sigma_{\text{gg}}(r_p)/\bar{\rho}_m$ . It is immediately clear that  $\Delta \Sigma_{\text{gg}}(r_p)$  can be obtained from the projected correlation function  $w_p(r_p)$ , which is routinely measured in large galaxy redshift surveys.

In terms of the classical 3D bias functions  $b_g^{3D}$ ,  $\mathcal{R}_{\text{gm}}^{3D}$ , and  $\Gamma_{\text{gm}}^{3D}$  (Cacciato et al. 2012), the galaxies can be unbiased with respect to the underlying dark matter distribution, if and only if the following conditions are true: they are not central galaxies, the occupation number of satellite galaxies obeys Poisson statistics ( $\beta = 1$ ), the normalized number density profile of satellite galaxies is identical to the one of the dark matter, and the occupational number of satellites is directly proportional to halo mass as  $\langle N_s \rangle = M\bar{n}_s/\rho$ . When central galaxies are added to the above conditions, one expects a strong scale dependence on small scales, due to the fact that central galaxies are strongly biased with respect to dark matter haloes. In the case of a non-Poissonian satellite distribution, one still expects  $b_g^{3D} = 1$  on large scales, but with a transition from 1 to  $\beta$ , roughly at the virial radius when moving towards the centre of the halo (see also fig. 3 in Cacciato et al. 2012). The same also holds for the case where the density profile of satellites follows that of dark matter (Cacciato et al. 2012).

Given all these reasons, as already pointed out by Cacciato et al. (2012), one expects scale independence on large scales (at a value dependent on halo model ingredients), with the transition to scale dependence on small scales (due to the effects of central galaxies) around the 1-halo to 2-halo transition. The same holds for the projected bias functions ( $b_g$ ,  $\mathcal{R}_{\text{gm}}$ , and  $\Gamma_{\text{gm}}$ ), which also carry a wealth of information regarding the non-linearity and stochasticity of halo occupation statistics, and consequently, galaxy formation.

This is demonstrated in Fig. 1 where we show the influence of different values of  $\sigma_c$ ,  $A_s$ ,  $\alpha_s$  and  $\beta$  on the bias function  $\Gamma_{\text{gm}}$  as a function of stellar mass. From the predictions one can clearly see how the different halo model ingredients influence the bias function. The halo model predicts, as mentioned before, scale independence above  $10 \text{ Mpc} h^{-1}$  and a significant scale dependence of galaxy bias on smaller scales, with the parameters  $\alpha_s$ ,  $A_s$ , and  $\beta$  having a significant influence at those scales. Any deviation from a pure Poissonian distribution of satellite galaxies will result in quite a significant feature at intermediate scales, therefore it would be a likely explanation for detected signs of stochasticity [as the deviation from unity will drive the stochasticity function  $\sigma_b$  or alternatively  $\epsilon$  away from 0,

as can be seen from equations (38)–(42)]. In Fig. 1, we also test the influence of having different  $\Omega_m$  and  $\sigma_8$  on the  $\Gamma_{\text{gm}}$  bias function, as generally, any bias function is a strong function of those two parameters (Dekel & Lahav 1999; Sheldon et al. 2004). We test this by picking four combinations of  $\Omega_m$  and  $\sigma_8$  drawn from the  $1\sigma$  confidence contours of Planck Collaboration (2016) measurements of the two parameters. Given the uncertainties of those parameters and their negligible influence on the  $\Gamma_{\text{gm}}$  bias function, the decision to fix the cosmology seems to be justified.

We would like to remind the reader that our implementation of the halo model does not include the scale dependence of the halo bias and the halo-exclusion (mutual exclusiveness of the spatial distribution of the haloes). Not including those effects can introduce errors on the 1-halo to 2-halo transition region that can be as large as 50 per cent (Cacciato et al. 2012; van den Bosch et al. 2013). However, the bias functions as defined using equations (47)–(49) are much more accurate and less susceptible to the uncertainties in the halo model, by being defined as ratios of the two-point correlation functions (Cacciato et al. 2012).

Despite of this, we decided to estimate the halo model parameters and the nature of galaxy bias using the fit to the  $\Delta \Sigma_{\text{gm}}(r_p)$  and  $w_p(r_p)$  signals separately, rather than the ratio of the two (using the  $\Gamma_{\text{gm}}$  bias function directly). This approach will still suffer from a possible bias due to the fact that we do not include the scale-dependent halo bias or the halo-exclusion in our model. This choice is motivated purely by the fact that the covariance matrix that would account for the cross-correlations between the lensing and clustering measurements cannot be properly taken into account when fitting the  $\Gamma_{\text{gm}}$  bias function directly. We investigate the possible bias in our results in Section 5.2.

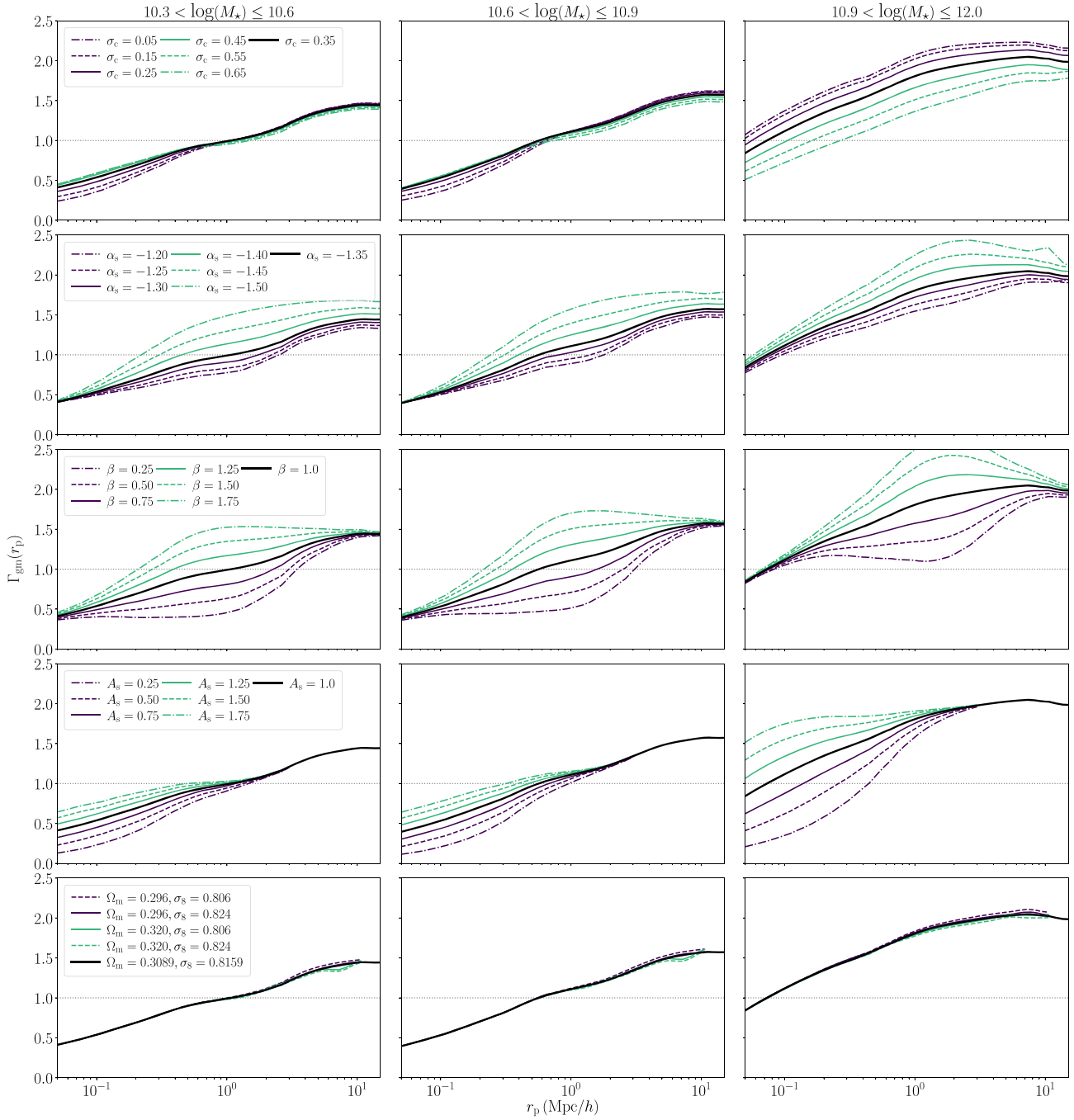
## 4 DATA AND SAMPLE SELECTION

### 4.1 Lens galaxy selection

The foreground galaxies used in this lensing analysis are taken from the Galaxy And Mass Assembly (hereafter GAMA) survey (Driver et al. 2011). GAMA is a spectroscopic survey carried out on the Anglo-Australian Telescope with the AAOmega spectrograph. Specifically, we use the information of GAMA galaxies from three equatorial regions, G9, G12, and G15 from GAMA II (Liske et al. 2015). We do not use the G02 and G23 regions, because the first one does not overlap with KiDS and the second one uses a different target selection compared to the one used in the equatorial regions. These equatorial regions encompass  $\sim 180 \text{ deg}^2$ , contain 180 960 galaxies (with  $nQ \geq 3$ , where the  $nQ$  is a measure of redshift quality) and are highly complete down to a Petrosian  $r$ -band magnitude  $r = 19.8$ . For the weak lensing measurements, we use all the galaxies in the three equatorial regions as potential lenses.

To measure their average lensing and projected clustering signals, we group GAMA galaxies in stellar mass bins, following previous lensing measurements by van Uitert et al. (2016) and Velliscig et al. (2017). The bin ranges were chosen this way to achieve a good signal-to-noise ratio in all bins and to measure the galaxy bias as a function of different stellar mass. The selection of galaxies can be seen in Fig. 2, and the properties we use in the halo model are shown in Table 1. Stellar masses are taken from version 19 of the stellar mass catalogue, an updated version of the catalogue created by Taylor et al. (2011), who fitted Bruzual & Charlot (2003) synthetic stellar population SEDs to the broad-band SDSS photometry assuming a Chabrier (2003) IMF and a Calzetti et al. (2000) dust law. The stellar masses in Taylor et al. (2011) agree well with

<sup>2</sup>In Dvornik et al. (2017), the same definition was used in all the calculations and plots shown, but erroneously documented in the paper. The equations (6) and (9) of that paper should have the same form as equations (50) and (54), as discussed in Appendix B.



**Figure 1.** Model predictions of scale dependence of the galaxy bias function  $\Gamma_{\text{gm}}$  (equation 49) for three stellar mass bins (defined in Table 1), with stellar masses given in units of  $[\log(M_*/[M_\odot/h^2])]$ . With the black solid line we show our fiducial halo model (with other parameters adapted from Cacciato et al. 2013), and the different green and violet lines show different values of  $\sigma_c$ ,  $\alpha_s$ ,  $\beta$ ,  $A_s$ , and combinations of  $\Omega_m$  and  $\sigma_8$ , row-wise, with values indicated in the legend. The full set of our fiducial parameters can be found in Table 2.

MagPhys-derived estimates, as shown by Wright et al. (2017). Despite the differences in the range of filters, star formation histories, obscuration laws, the two estimates agree within 0.2 dex for 95 per cent of the sample.

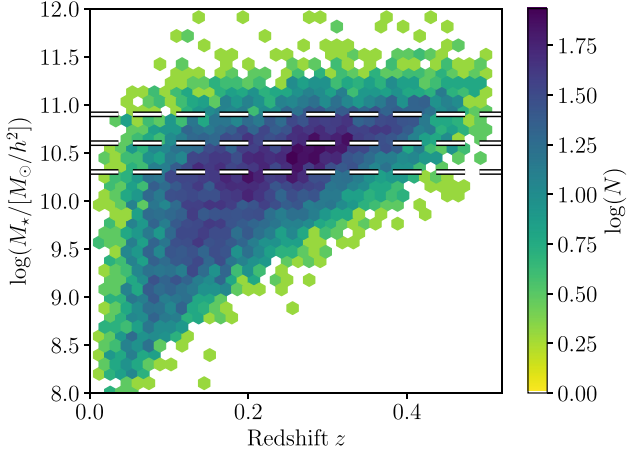
#### 4.2 Measurement of the $\Delta\Sigma_{\text{gm}}(r_p)$ signal

We use imaging data from 180 deg<sup>2</sup> of KiDS (de Jong et al. 2015; Kuijken et al. 2015) that overlaps with the GAMA survey (Driver

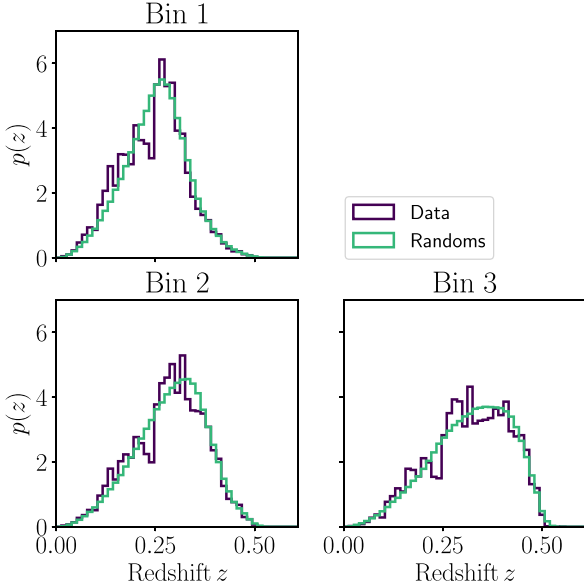
et al. 2011) to obtain shape measurements of background galaxies. KiDS is a four-band imaging survey conducted with the OmegaCAM CCD mosaic camera mounted at the Cassegrain focus of the VLT Survey Telescope (VST); the camera and telescope combination provide us with a fairly uniform point spread function (PSF) across the field-of-view.

We use shape measurements based on the *r*-band images, which have an average seeing of 0.66 arcsec. The image reduction, photometric redshift calibration, and shape measure-





**Figure 2.** Stellar mass versus redshift of galaxies in the GAMA survey that overlap with KiDS. The full sample is shown with hexagonal density plot and the dashed lines show the cuts for the three stellar mass bins used in our analysis.



**Figure 3.** A comparison between the redshift distribution of galaxies in the data and the matched galaxies in GAMA random catalogue (Farrow et al. 2015) for our three stellar mass bins. We use the same set of randoms for both galaxy clustering and galaxy–galaxy lensing measurements.

ment analysis is described in detail in Hildebrandt et al. (2017).

We measure galaxy shapes using calibrated *lensfit* shape catalogues (Miller et al. 2013; see also Fenech Conti et al. 2017 where the calibration methodology is described), which provides galaxy ellipticities ( $\epsilon_1, \epsilon_2$ ) with respect to an equatorial coordinate system. For each source–lens pair, we compute the tangential  $\epsilon_t$  and cross component  $\epsilon_\times$  of the source’s ellipticity around the position of the lens:

$$\begin{bmatrix} \epsilon_t \\ \epsilon_\times \end{bmatrix} = \begin{bmatrix} -\cos(2\phi) & -\sin(2\phi) \\ \sin(2\phi) & -\cos(2\phi) \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix}, \quad (52)$$

where  $\phi$  is the angle between the  $x$ -axis and the lens–source separation vector.

The azimuthal average of the tangential ellipticity of a large number of galaxies in the same area of the sky is an unbiased estimate of the shear. On the other hand, the azimuthal average of the cross ellipticity over many sources is unaffected by gravitational lensing and should average to zero (Schneider 2003). Therefore, the cross ellipticity is commonly used as an estimator of possible systematics in the measurements such as non-perfect PSF deconvolution, centroid bias, and pixel level detector effects (Mandelbaum 2017). Each lens–source pair is then assigned a weight

$$\tilde{w}_{ls} = w_s \left( \tilde{\Sigma}_{cr,ls}^{-1} \right)^2, \quad (53)$$

which is the product of the *lensfit* weight  $w_s$  assigned to the given source ellipticity and the square of  $\tilde{\Sigma}_{cr,ls}^{-1}$  – the effective inverse critical surface mass density, which is a geometric term that down-weights lens–source pairs that are close in redshift. We compute the effective inverse critical surface mass density for each lens using the spectroscopic redshift of the lens  $z_1$  and the full normalized redshift probability density of the sources,  $n(z_s)$ , calculated using the direct calibration method presented in Hildebrandt et al. (2017).

The effective inverse critical surface density can be written as

$$\tilde{\Sigma}_{cr,ls}^{-1} = \frac{4\pi G}{c^2} (1 + z_1)^2 D(z_1) \int_{z_1}^{\infty} \frac{D(z_1, z_s)}{D(z_s)} n(z_s) dz_s. \quad (54)$$

The galaxy source sample is specific to each lens redshift with a minimum photometric redshift  $z_s = z_1 + \delta_z$ , with  $\delta_z = 0.2$ , where  $\delta_z$  is an offset to mitigate the effects of contamination from the group galaxies (for details see also the methods section and appendix of Dvornik et al. 2017). We determine the source redshift distribution  $n(z_s)$  for each sample, by applying the sample photometric redshift selection to a spectroscopic catalogue that has been weighted to reproduce the correct galaxy colour-distributions in KiDS (for details see Hildebrandt et al. 2017). Thus, the ESD can be directly computed in bins of projected distance  $r_p$  to the lenses as

$$\Delta \Sigma_{gm}(r_p) = \left[ \frac{\sum_{ls} \tilde{w}_{ls} \epsilon_{t,s} \Sigma'_{cr,ls}}{\sum_{ls} \tilde{w}_{ls}} \right] \frac{1}{1 + \bar{m}}, \quad (55)$$

where  $\Sigma'_{cr,ls} \equiv 1/\tilde{\Sigma}_{cr,ls}^{-1}$  and the sum is over all source–lens pairs in the distance bin, and

$$\bar{m} = \frac{\sum_i w'_i m_i}{\sum_i w'_i}, \quad (56)$$

is an average correction to the ESD profile that has to be applied to correct for the multiplicative bias  $m$  in the *lensfit* shear estimates. The sum goes over thin redshift slices for which  $m$  is obtained using the method presented in Fenech Conti et al. (2017), weighted by  $w' = w_s D(z_1, z_s)/D(z_s)$  for a given lens–source sample. The value of  $\bar{m}$  is around  $-0.014$ , independent of the scale at which it is computed. Furthermore, we subtract the signal around random points using the random catalogues from Farrow et al. (2015) (for details see analysis in the appendix of Dvornik et al. 2017).

#### 4.3 Measurement of the $w_p(r_p)$ profile

We compute the 3D autocorrelation function of our three lens samples using the Landy & Szalay (1993) estimator. For this we use the same random catalogue and procedure as described in Farrow et al. (2015), applicable to the GAMA data. To minimize the effect of redshift-space distortions in our analysis, we project the 3D autocorrelation function along the line of sight:

$$w_p(r_p) = 2 \int_0^{\Pi_{\max}=100 \text{ Mpc}/h} \xi(r_p, \Pi) d\Pi. \quad (57)$$

For practical reasons, the above integral is evaluated numerically. This calls for consideration of our integration limits, particularly the choice of  $\Pi_{\max}$ . Theoretically one would like to integrate out to infinity in order to completely remove the effect of redshift space distortions and to encompass the full clustering signal on large scales. We settle for  $\Pi_{\max} = 100 \text{ Mpc}/h$ , in order to project the correlation function on the separations we are interested in (with a maximum  $r_p = 10 \text{ Mpc}/h$ ). We use the publicly available code SWOT<sup>3</sup> (Coupon et al. 2012) to compute  $\xi(r_p, \Pi)$  and  $w_p(r_p)$ , and to get bootstrap estimates of the covariance matrix on small scales. The code was tested against results from Farrow et al. (2015) using the same sample of galaxies and updated random catalogues (internal version 0.3), reproducing the results in detail. Randoms generated by Farrow et al. (2015) contain around 750 times more galaxies than those in GAMA samples. Fig. 3 shows the good agreement between the redshift distributions of the GAMA galaxies and the random catalogues for the three stellar mass bins.

The clustering signal  $w_p(r_p)$  as well as the lensing signal  $\Delta\Sigma_{\text{gm}}(r_p)$  are shown in Fig. 4, in the right- and left-hand panel, respectively. They are shown together with MCMC best-fitting profiles as described in Section 4.5, using the halo model as described in Section 3. The best fit is a single model used for all stellar masses and not independent for the three bins we are using. In order to obtain the galaxy bias function  $\Gamma_{\text{gm}}(r_p)$  (equation 49), we project the clustering signal according to equation (51). The plot of this resulting function can be seen in Fig. 5.

#### 4.4 Covariance matrix estimation

Statistical error estimates on the lensing signal and projected galaxy clustering signal are obtained using an analytical covariance matrix. As shown in Dvornik et al. (2017), estimating the covariance matrix from data can become challenging given the small number of independent data patches in GAMA. This becomes even more challenging when one wants to include in the mixture the covariance for the projected galaxy clustering and all the possible cross terms between the two. The analytical covariance matrix we use is composed of three main parts: a Gaussian term, non-Gaussian term, and the super-sample covariance (SSC), which accounts for all the modes outside of our KiDS x GAMA survey window. It is based on previous work by Takada & Jain (2009), Joachimi, Schneider & Eifler (2008), Pielorz et al. (2010), Takada & Hu (2013), Li, Hu & Takada (2014), Marian, Smith & Angulo (2015), Singh et al. (2017), and Krause & Eifler (2017), and extended to support multiple lens bins and cross terms between lensing and projected galaxy clustering signals. The covariance matrix was tested against published results in these individual papers, as well as against real data estimates on small scales and mocks as used by van Uitert et al. (2018). Further details and terms used can be found in Appendix A. We first evaluate our covariance matrix for a set of fiducial model parameters and use this in our MCMC fit and then take the best-fitting values and re-evaluate the covariance matrix for the new best-fitting halo model parameters. After carrying out the re-fitting procedure, we find out that the updated covariance matrix and halo model parameters do not affect the results of our fit, and thus the original estimate of the covariance matrix is appropriate to use throughout the analysis.

#### 4.5 Fitting procedure

The free parameters for our model are listed in Table 2, together with their fiducial values. We use a Bayesian inference method in order to obtain full posterior probabilities using a Monte Carlo Markov Chain (MCMC) technique; more specifically we use the emcee Python package (Foreman-Mackey et al. 2013). The likelihood  $\mathcal{L}$  is given by

$$\mathcal{L} \propto \exp \left[ -\frac{1}{2} (\mathbf{O}_i - \mathbf{M}_i)^T \mathbf{C}_{ij}^{-1} (\mathbf{O}_j - \mathbf{M}_j) \right], \quad (58)$$

where  $\mathbf{O}_i$  and  $\mathbf{M}_i$  are the measurements and model predictions in radial bin  $i$ , and  $\mathbf{C}_{ij}^{-1}$  is the element of the inverse covariance matrix that accounts for the correlation between radial bins  $i$  and  $j$ . In the fitting procedure we use the inverse covariance matrix as described in Section 4.4 and Appendix A. We use wide flat priors for all the parameters (given in Table 2). The halo model (halo mass function and the power spectrum) is evaluated at the median redshift for each sample.

We run the sampler using 120 walkers, each with 12 000 steps (for a combined number of 14 400 000 samples), out of which we discard the first 1000 burn-in steps (120 000 samples). The resulting MCMC chains are well converged according to the integrated autocorrelation time test.

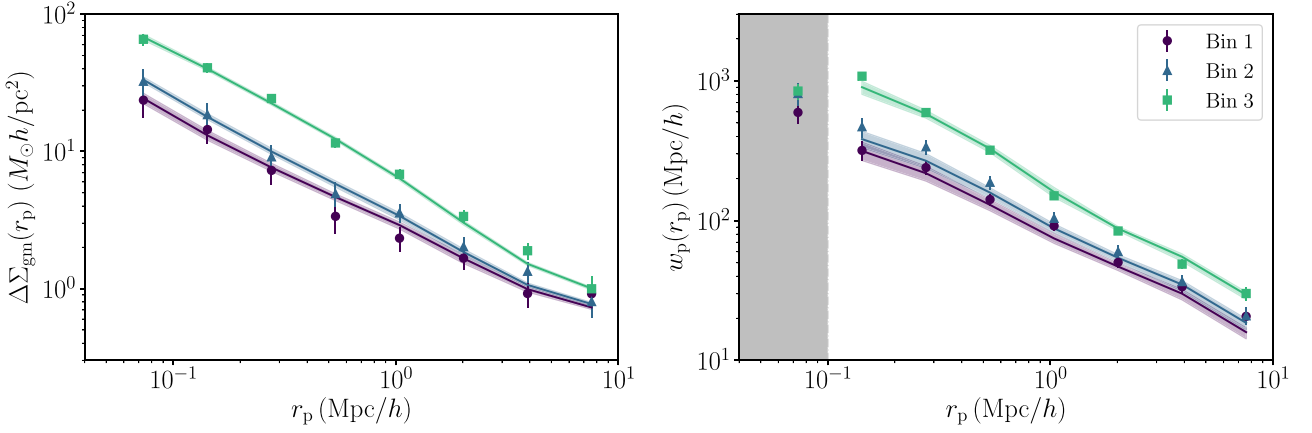
## 5 RESULTS

### 5.1 KiDS and GAMA results

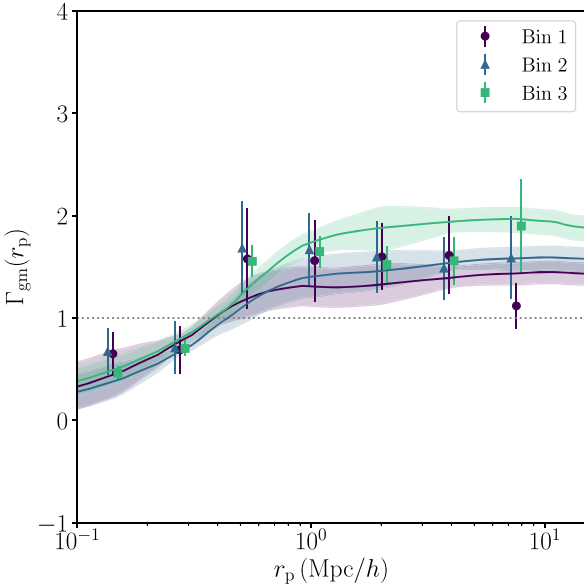
We fit the halo model as described in Section 4.5 to the measured projected galaxy clustering signal  $w_p(r_p)$  and the galaxy–galaxy lensing signal  $\Delta\Sigma_{\text{gm}}(r_p)$ , using the covariance matrix as described in Section 4.4. The resulting best fits are presented in Fig. 4 (together with the measurements and their respective  $1\sigma$  errors obtained by taking the square root of the diagonal elements of the analytical covariance matrix). The measured halo model parameters, together with the  $1\sigma$  uncertainties are summarized in Table 2. Their full posterior distributions are shown in Fig. B1. The fit of our halo model to both the galaxy–galaxy lensing signal and projected galaxy clustering signal, using the full covariance matrix accounting for all the possible cross-correlations, has a reduced  $\chi^2_{\text{red}} (\equiv \chi^2/\text{d.o.f.})$  equal to 1.15, which is an appropriate fit, given the 33 degrees of freedom (d.o.f.). We urge readers not to rely on the ‘chi-by-eye’ in Figs 4 and 5 due to highly correlated data points (the correlations of which can be seen in Fig. A1) and the joint fit of the halo model to the data.

Due to the fact that we are only using samples with relatively high stellar masses, we are unable to sample the low-mass portion of the stellar mass function, evident in our inability to properly constrain the  $\gamma_1$  parameter, which describes the behaviour of the stellar mass function at low halo mass. Mostly because of this, our results for the HOD parameters are different compared to those obtained by van Uitert et al. (2016), who analysed the full GAMA sample. There is also a possible difference arising due to the available overlap of KiDS and GAMA surveys used in van Uitert et al. (2016) and our analysis, as van Uitert et al. (2016) used the lensing data from only  $100 \text{ deg}^2$  of the KiDS data, released before the shear catalogues used by Hildebrandt et al. (2017) and Dvornik et al. (2017), amongst others, became available. Our inferred HOD parameters are also in broad agreement with the ones obtained by Cacciato, van Uitert & Hoekstra (2014) for a sample of SDSS galaxies.

<sup>3</sup><http://jeancoupon.com/swot>



**Figure 4.** The stacked ESD profile (*left-hand panel*) and projected galaxy clustering signal (*right-hand panel*) of the three stellar mass bins in the GAMA galaxy sample defined in Table 1. The solid lines represent the best-fitting halo model as obtained using an MCMC fit, with the 68 per cent confidence interval indicated with a shaded region. Using those two measurements we obtain the bias function  $\Gamma_{\text{gm}}(r_p)$ . We do not use the measurements in the grey band in our fit, as the clustering measurements are affected by blending in this region. The best-fitting halo model parameters are listed in Table 2.



**Figure 5.** The  $\Gamma_{\text{gm}}(r_p)$  bias function as measured using a combination of projected galaxy clustering and galaxy–galaxy lensing signals, shown for the three stellar mass bins as used throughout this paper. The solid lines represent the best-fitting halo model as obtained using an MCMC fit to the projected galaxy clustering and galaxy–galaxy lensing signal, combined to obtain  $\Gamma_{\text{gm}}(r_p)$ , as described in Section 3. The colour bands show the 68 per cent confidence interval propagated from the best-fitting model. Error bars on the data are obtained by propagating the appropriate sub-diagonals of the covariance matrix and thus do not show the correct correlations between the data points and also overestimate the sample variance and super-sample covariance contributions.

The main result of this work is the  $\Gamma_{\text{gm}}(r_p)$  bias function, presented in Fig. 5, together with the best-fitting MCMC result – obtained by projecting the measured galaxy clustering result according to equation (51) – and combining with the galaxy–galaxy lensing result according to equation (49). The obtained  $\Gamma_{\text{gm}}(r_p)$  bias function from the fit is scale dependent, showing a clear transition around  $2 \text{ Mpc } h^{-1}$ , in the 1-halo to 2-halo regime, where the function slowly transitions towards a constant value on even larger scales, beyond the range studied here (as predicted in Cacciato et al. 2012). Given

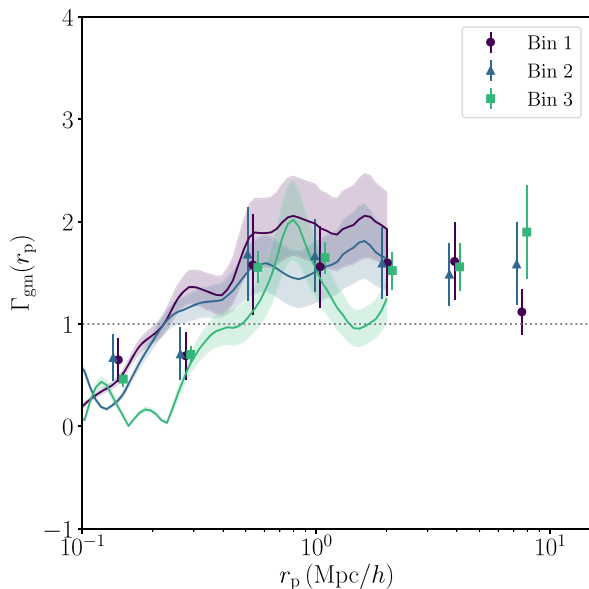
the parameters obtained using the halo model fit to the data, the preferred value of  $\beta$  is larger than unity with  $\beta = 1.67^{+0.15}_{-0.16}$ , which indicates that the satellite galaxies follow a super-Poissonian distribution inside their host dark matter haloes, and are thus responsible for the deviations from constant in our  $\Gamma_{\text{gm}}(r_p)$  bias function at intermediate scales. Following the formulation by Cacciato et al. (2012), this also means that the galaxy bias, as measured, is highly non-deterministic. As seen by the predictions shown in Fig. 1, the deviation of  $\beta$  from unity alone is not sufficient to explain the full observed scale dependence of the  $\Gamma_{\text{gm}}(r_p)$  bias function. Given the best-fitting parameter values using the MCMC fit of the halo model, the non-unity of the mass-concentration relation normalization  $A_s$  and other CSMF parameters (but most importantly the  $\alpha_s$  parameter, which governs the power-law behaviour of the satellite CSMF) are also responsible for the total contribution to the observed scale dependence, and thus the stochastic behaviour of the galaxy bias on all scales observed.

## 5.2 Investigation of the possible bias in the results

Due to the fact that we have decided to fit the model to the  $\Delta\Sigma_{\text{gm}}(r_p)$  and  $w_p(r_p)$  signals, we investigate how this choice might have biased our results. To check this we repeat our analysis using the  $\Gamma_{\text{gm}}(r_p)$  bias function directly. As our data vector we take the ratio of the projected signals as shown in Fig. 5 and we use the appropriately propagated sub-diagonals of the covariance matrix as a rough estimate of the total covariance matrix. Such a covariance matrix does not show the correct correlations between the data points (and the bins) and also overestimates the sample variance and super-sample covariance contributions. Never the less the ratio of the diagonals as an estimate of the errors is somewhat representative of the errors on the measured  $\Gamma_{\text{gm}}(r_p)$  bias function. The fit procedure (except for a different data vector, covariance, and output of the model) follows the method presented in Section 4.5. Using this, we obtain the best-fitting values that are shown in Fig. B1, marked with blue points and lines, together with the full posterior distributions from the initial fit. The resulting fit has a  $\chi^2_{\text{red}}$  equal to 1.29, with 9 degrees of freedom. As the results are consistent with the results that we obtain using a fit to the  $\Delta\Sigma_{\text{gm}}(r_p)$  and  $w_p(r_p)$  signals separately, it seems that, at least for this study, the halo model as described does not bias the overall conclusions of our analysis.

**Table 2.** Summary of the lensing results obtained using MCMC halo model fit to the data. Here,  $M_0$  is the normalization of the stellar to halo mass relation,  $M_1$  is the characteristic mass scale of the same stellar to halo mass relation,  $A_c$  is the normalization of the concentration-mass relation,  $\sigma_c$  is the scatter between the stellar and halo mass,  $\gamma_1$  and  $\gamma_2$  are the low and high-mass slopes of the stellar to halo mass relation,  $A_s$  is the normalization of the concentration-mass relation for satellite galaxies,  $\alpha_s$ ,  $b_0$ , and  $b_1$  govern the behaviour of the CSMF of satellite galaxies, and  $\beta$  is the Poisson parameter. All parameters are defined in Section 3, using equations (29)–(42).

	$\log(M_0/[M_\odot/h^2])$	$\log(M_1/[M_\odot/h])$	$A_c$	$\sigma_c$	$\gamma_1$	$\gamma_2$
Fiducial	9.6	11.25	1.0	0.35	3.41	0.99
Priors	[7.0, 13.0]	[9.0, 14.0]	[0.0, 5.0]	[0.05, 2.0]	[0.0, 10.0]	[0.0, 10.0]
Posteriors	$8.75^{+1.62}_{-1.28}$	$11.13^{+1.10}_{-1.11}$	$1.33^{+0.20}_{-0.19}$	$0.25^{+0.24}_{-0.18}$	$2.16^{+4.43}_{-1.52}$	$1.32^{+0.51}_{-0.34}$
	$A_s$	$\alpha_s$	$b_0$	$b_1$	$\beta$	
Fiducial	1.0	−1.34	−1.15	0.59	1.0	
Priors	[0.0, 5.0]	[−5.0, 5.0]	[−5.0, 5.0]	[−5.0, 5.0]	[0.0, 2.0]	
Posteriors	$0.24^{+0.30}_{-0.14}$	$−1.36^{+0.19}_{-0.13}$	$−0.71^{+0.34}_{-0.55}$	$0.13^{+0.29}_{-0.30}$	$1.67^{+0.15}_{-0.16}$	



**Figure 6.** The  $\Gamma_{\text{gm}}(r_p)$  bias function as measured using the combination of projected galaxy clustering and galaxy–galaxy lensing signals, shown for the three stellar mass bins as used throughout this paper. The solid lines represent the same measurement repeated on the EAGLE simulation, with the colour bands showing the  $1\sigma$  errors. Note that those measurements are noisy due to the fact that the EAGLE simulation box is rather small, resulting in a relatively low number of galaxies in each bin (factor of around 26 lower, compared to the data). Due to the box size, we can also only show the measurement to about  $2 \text{ Mpc } h^{-1}$ .

### 5.3 Comparison with EAGLE simulation

In Fig. 6, we compare our measurements of the GAMA and KiDS data to the same measurements made using the hydrodynamical EAGLE simulation (Schaye et al. 2015; McAlpine et al. 2016). EAGLE consists of state-of-the-art hydrodynamical simulations, including sub-grid interaction mechanisms between stellar and galactic energy sources. EAGLE is optimized such that the simulations reproduce a universe with the same stellar mass function as our own (Schaye et al. 2015). We follow the same procedure as with the data, by separately measuring the projected galaxy clustering signal and the galaxy–galaxy lensing signal and later combining the two accordingly. We measure the 3D galaxy clustering using the Landy & Szalay (1993) estimator, closely following the procedure outlined

in Artale et al. (2017). We adopt the same  $\Pi_{\text{max}} = 34 \text{ Mpc } h^{-1}$  as used by Artale et al. (2017) in order to project the 3D galaxy clustering  $\xi(r_p, \Pi)$  to  $w_p(r_p)$ , which represents  $\sim L/2$  of the EAGLE box (Artale et al. 2017); see also equation (51). This limits the EAGLE measurements to a maximum scales of  $r_p < 2 \text{ Mpc } h^{-1}$ . As we do not require an accurate covariance matrix for the EAGLE results (we do not fit any model to it), we adopt a Jackknife covariance estimator using eight equally sized sub-volumes. The measured EAGLE projected galaxy clustering signal is in good agreement with the GAMA measurements in detail, a result also found in Artale et al. (2017).

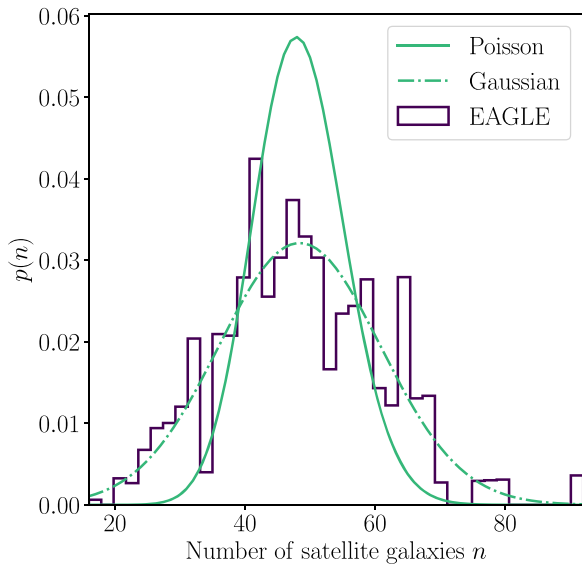
To estimate the galaxy–galaxy lensing signal of galaxies in EAGLE, we use the excess surface density (i.e. lensing signal) of galaxies in EAGLE calculated by Velliscig et al. (2017). We again select the galaxies in the three stellar mass bins, but in order to mimic the magnitude-limited sample we have adopted in our measurements of the galaxy–galaxy lensing signal on GAMA and KiDS, we have to weight our galaxies in the selection according to the satellite fraction as presented in Velliscig et al. (2017).

Our two measurements (projected galaxy clustering and the galaxy–galaxy lensing) are then combined according to the definition of the  $\Gamma_{\text{gm}}(r_p)$  bias function, which is shown in Fig. 6. There we directly compare the bias function as measured in the KiDS and GAMA data to the one obtained from the EAGLE hydrodynamical simulation (shown with full lines). The results from EAGLE are noisy, due to the fact that one is limited by the number of galaxies present in EAGLE.

Using the EAGLE simulations, we can directly access the properties of the satellite galaxies residing in the main haloes present in the simulation. We select a narrow bin in halo masses of groups present in the simulation (between 12.0 and 12.2 in  $\log(M/M_\odot)$ ) and count the number of sub-haloes (galaxies). The resulting histogram, showing the relative abundance of satellite galaxies can be seen in Fig. 7. We also show the Poisson distribution with the same mean as the EAGLE data, as well as the Gaussian distribution with the same mean and standard deviation as the distribution of the satellite galaxies in our sample. It can be immediately seen that the distribution of satellite galaxies at a fixed halo mass does not follow a Poisson distribution, and it is significantly wider (thus indeed being super-Poissonian).

The comparison never the less shows that the galaxy bias is intrinsically scale dependent, and the shape of it suggests that it can be attributed to the non-Poissonian behaviour of satellite galaxies (and to lesser extent also to the precise distribution of satellites in the dark matter halo, governed by  $\alpha_s$  and  $A_s$  in the halo model).





**Figure 7.** Distribution of satellite galaxies in a halo of fixed mass within  $12.0 < \log(M/M_\odot) < 12.2$  (histogram). This can be compared to a Poisson distribution with the same mean (solid curve) and a Gaussian distribution with the same mean and standard deviation as the data (dot-dashed curve).

## 6 DISCUSSION AND CONCLUSIONS

We have measured the projected galaxy clustering signal and galaxy–galaxy lensing signal for a sample of GAMA galaxies as a function of their stellar mass. In this analysis, we use the KiDS data covering  $180 \text{ deg}^2$  of the sky (Hildebrandt et al. 2017), that fully overlaps with the three equatorial patches from the GAMA survey that we use to determine three stellar mass-selected lens galaxy samples. We have combined our results to obtain the  $\Gamma_{\text{gm}}(r_p)$  bias function in order to unveil the hidden factors and origin of galaxy biasing in light of halo occupation models and the halo model, as presented in the theoretical work of Cacciato et al. (2012). We have used that formalism to fit to the data to constrain the parameters that contribute to the observed scale dependence of the galaxy bias, and see which parameters exactly carry information about the stochasticity and non-linearity of the galaxy bias, as observed. Due to the limited area covered by the both surveys, the covariance matrix used in this analysis was estimated using an analytical prescription, for which details can be found in Appendix A.

Our results show a clear trend that galaxy bias cannot be simply treated with a linear and/or deterministic approach. We find that the galaxy bias is inherently stochastic and non-linear due to the fact that satellite galaxies do not strictly follow a Poissonian distribution and that the spatial distribution of satellite galaxies also does not follow the NFW profile of the host dark matter halo. The main origin of the non-linearity of galaxy bias can be attributed to the fact that the central galaxy itself is heavily biased with respect to the dark matter halo in which it is residing. Those findings give additional support for the predictions presented by Cacciato et al. (2012), as their conclusions, based only on some fiducial model, are in line with our finding for a real subset of galaxies. We observe the same trends in the cosmological hydrodynamical simulation EAGLE, albeit out to smaller scales. We have also shown that the  $\Gamma_{\text{gm}}(r_p)$  bias function can, by itself, measure the properties of galaxy bias that would otherwise require the full knowledge of the  $b_g(r_p)$  and  $\mathcal{R}_{\text{gm}}(r_p)$  bias functions.

Our results are also in a broad agreement with recent findings of Gruen et al. (2017) and Friedrich et al. (2017), who used the density split statistics to measure the cosmological parameters in SDSS (Rozo et al. 2015) and DES (Drlica-Wagner et al. 2018) data, and as a byproduct, also the  $b$  and  $r$  functions directly (at angular scales around  $20 \text{ arcmin}$ , which correspond to  $3.5\text{--}7 \text{ Mpc h}^{-1}$  at redshifts of  $0.2\text{--}4.5$ ). They find that the SDSS and DES data strongly prefer a stochastic bias with super-Poissonian behaviour. To obtain an independent measurement of galaxy bias and to further confirm our results, we could use this method on our selection of galaxies, as well as the reconstruction method of Simon & Hilbert (2018). This work is, however, out of the scope of this paper.

Our findings show a remarkable wealth of information that halo occupation models are carrying in regard of understanding the nature of galaxy bias and its influence on cosmological analyses using the combination of galaxy–galaxy lensing and galaxy clustering. These results also show that the theoretical framework, as presented by Cacciato et al. (2012), is able to translate the constraints on galaxy biasing into constraints on galaxy formation and measurements of cosmological parameters.

As an extension of this work, we could fold in the cosmic shear measurements of the same sample of galaxies, and thus constrain the galaxy bias and the sources of non-linearity and stochasticity further. This would allow a direct measurement of all three bias functions [ $\Gamma_{\text{gm}}(r_p)$ ,  $b_g(r_p)$ , and  $\mathcal{R}_{\text{gm}}(r_p)$ ], which could then be used directly in cosmological analyses. On the other hand, for a more detailed study of the HOD beyond those parameters that influence the galaxy bias, we could include the stellar mass (or luminosity) function in the joint fit. We leave such exercises open for future studies.

## ACKNOWLEDGEMENTS

We thank the anonymous referee for their very useful comments and suggestions. AD would like to thank Marcello Cacciato for all the useful discussions, support, and the hand-written notes provided on the finer aspects of the theory used in this paper.

KK acknowledges support by the Alexander von Humboldt Foundation. HHo acknowledges support from Vici grant 639.043.512, financed by the Netherlands Organisation for Scientific Research (NWO). This work is supported by the Deutsche Forschungsgemeinschaft in the framework of the TR33 ‘The Dark Universe’. CH acknowledges support from the European Research Council under grant number 647112. HHi is supported by an Emmy Noether grant (No. Hi 1495/2-1) of the Deutsche Forschungsgemeinschaft. AA is supported by a LSSTC Data Science Fellowship. RN acknowledges support from the German Federal Ministry for Economic Affairs and Energy (BMWi) provided via DLR under project no. 50QE1103.

This research is based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under programme IDs 177.A-3016, 177.A-3017, and 177.A-3018, and on data products produced by Target/OmegaCEN, INAF-OACN, INAF-OAPD, and the KiDS production team, on behalf of the KiDS consortium.

GAMA is a joint European-Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalogue is based on data taken from the Sloan Digital Sky Survey and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programmes including GALEX MIS, VST KiDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT, and

ASKAP providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO, and the participating institutions. The GAMA website is <http://www.gama-survey.org>.

This work has made use of PYTHON (<http://www.python.org>), including the packages NUMPY (<http://www.numpy.org>) and SCIPY (<http://www.scipy.org>). The halo model is built upon the hmf PYTHON package by Murray, Power & Robotham (2013). Plots have been produced with MATPLOTLIB (Hunter 2007) and CORNER.PY (Foreman-Mackey 2016).

*Author contributions:* All authors contributed to writing and development of this paper. The authorship list reflects the lead authors (AD, HH, KK, PS) followed by two alphabetical groups. The first alphabetical group includes those who are key contributors to both the scientific analysis and the data products. The second group covers those who have made a significant contribution either to the data products or to the scientific analysis.

## REFERENCES

- Amon A., et al., 2018a, *MNRAS*, sty1624,  
Amon A. et al., 2018b, *MNRAS*, 477, 4285  
Artale M. C. et al., 2017, *MNRAS*, 470, 1771  
Bartelmann M., Schneider P., 2001, *Phys. Rep.*, 340, 291  
Brouwer M. M. et al., 2016, *MNRAS*, 462, 4451  
Brouwer M. M. et al., 2017, *MNRAS*, 466, 2547  
Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000  
Buddendiek A. et al., 2016, *MNRAS*, 456, 3886  
Cacciato M., van den Bosch F. C., More S., Li R., Mo H. J., Yang X., 2009, *MNRAS*, 394, 929  
Cacciato M., Lahav O., van den Bosch F. C., Hoekstra H., Dekel A., 2012, *MNRAS*, 426, 566  
Cacciato M., van den Bosch F. C., More S., Mo H., Yang X., 2013, *MNRAS*, 430, 767  
Cacciato M., van Uitert E., Hoekstra H., 2014, *MNRAS*, 437, 377  
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, 533, 682  
Chabrier G., 2003, *ApJ*, 586, L133  
Cooray A., Sheth R., 2002, *Phys. Rep.*, 372, 1  
Coupon J. et al., 2012, *A&A*, 542, A5  
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371  
de Jong J. T. A. et al., 2015, *A&A*, 582, A62  
de la Torre S. et al., 2017, *A&A*, 608, A44  
Dekel A., Lahav O., 1999, *ApJ*, 520, 24  
Dekel A., Rees M. J., 1987, *Nature*, 326, 455  
Driver S. P. et al., 2011, *MNRAS*, 413, 971  
Drlica-Wagner A. et al., 2018, *ApJS*, 235, 33  
Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, *MNRAS*, 390, L64  
Dvornik A. et al., 2017, *MNRAS*, 468, 3251  
Eisenstein D. J., Hu W., 1998, *ApJ*, 496, 605  
Farrow D. J. et al., 2015, *MNRAS*, 454, 2120  
Fenech Conti I., Herbonnet R., Hoekstra H., Merten J., Miller L., Viola M., 2017, *MNRAS*, 467, 1627  
Foreman-Mackey D., 2016, *J. Open Source Softw.*, 2, 1  
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306  
Friedrich O. et al., 2017, preprint ([arXiv:1710.05162](https://arxiv.org/abs/1710.05162))  
Gruen D. et al., 2017, preprint ([arXiv:1710.05045](https://arxiv.org/abs/1710.05045))  
Hilbert S., Hartlap J., White S. D. M., Schneider P., 2009, *A&A*, 499, 31  
Hildebrandt H. et al., 2017, *MNRAS*, 465, 1454  
Hoekstra H., van Waerbeke L., Gladders M. D., Mellier Y., Yee H. K. C., 2002, *ApJ*, 577, 604  
Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90  
Joachimi B., Schneider P., Eifler T., 2008, *A&A*, 477, 43  
Jullo E. et al., 2012, *ApJ*, 750, 37  
Kaiser N., 1984, *ApJ*, 284, L9  
Krause E., Eifler T., 2017, *MNRAS*, 470, 2100  
Krause E., Hirata C. M., 2010, *A&A*, 523, A28  
Kuijken K. et al., 2015, *MNRAS*, 454, 3500  
Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64  
Li Y., Hu W., Takada M., 2014, *Phys. Rev. D*, 89, 1  
Liske J. et al., 2015, *MNRAS*, 452, 2087  
Mandelbaum R., 2017, preprint ([arXiv:1710.03235](https://arxiv.org/abs/1710.03235))  
Mandelbaum R., Seljak U., Baldauf T., Smith R. E., 2010, *MNRAS*, 405, 2078  
Marian L., Smith R. E., Angulo R. E., 2015, *MNRAS*, 451, 1418  
McAlpine S. et al., 2016, *A&C*, 15, 72  
Mead A. J., Peacock J. A., Heymans C., Joudaki S., Heavens A. F., 2015, *MNRAS*, 454, 1958  
Miller L. et al., 2013, *MNRAS*, 429, 2858  
Mo H., van den Bosch F., White S., 2010, *Galaxy Formation and Evolution*. Cambridge Univ. Press, Cambridge  
Murray S., Power C., Robotham A., 2013, *A&C*, 3-4, 23  
Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493  
Peacock J. A., Smith R. E., 2000, *MNRAS*, 318, 1144  
Pielorz J., Rödiger J., Tereno I., Schneider P., 2010, *A&A*, 514, A79  
Planck Collaboration, 2016, *A&A*, 594, A1  
Rozo E. et al., 2015, *MNRAS*, 461, 1431  
Schaye J. et al., 2015, *MNRAS*, 446, 521  
Schneider P., 2003, preprint ([arXiv:0306465](https://arxiv.org/abs/0306465))  
Schneider P., 2006, in Meylan G., Jetzer P., North P., Schneider P., Kochanek C., Wambsgans J., eds, *Gravitational Lensing: Strong, Weak and Micro – Saas-Fee Advanced Course 33*, Springer-Verlag, Berlin Heidelberg, p. 269  
Seitz S., Schneider P., Ehlers J., 1994, *Class. Quantum Gravity*, 11, 2345  
Seljak U., 2000, *MNRAS*, 318, 203  
Sheldon E. S. et al., 2004, *AJ*, 127, 2544  
Sifón C. et al., 2015, *MNRAS*, 454, 3938  
Simon P., Hilbert S., 2018, *A&A*, 613, A15  
Simon P., Hetterscheidt M., Schirmer M., Erben T., Schneider P., Wolf C., Meisenheimer K., 2007, *A&A*, 461, 861  
Singh S., Mandelbaum R., Seljak U., Slosar A., Vazquez Gonzalez J., 2017, *MNRAS*, 471, 3827  
Takada M., Hu W., 2013, *Phys. Rev. D*, 87, 1  
Takada M., Jain B., 2009, *MNRAS*, 395, 2065  
Taylor E. N. et al., 2011, *MNRAS*, 418, 1587  
Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, *ApJ*, 724, 878  
van den Bosch F. C., More S., Cacciato M., Mo H., Yang X., 2013, *MNRAS*, 430, 725  
van Uitert E. et al., 2016, *MNRAS*, 459, 3251  
van Uitert E. et al., 2018, *MNRAS*, 476, 4662  
Velliscig M. et al., 2017, *MNRAS*, 471, 2856  
Viola M. et al., 2015, *MNRAS*, 452, 3529  
Wang L. et al., 2013, *MNRAS*, 431, 648  
Wang Y., Yang X., Mo H. J., van den Bosch F. C., Weinmann S. M., Chu Y., 2008, *ApJ*, 687, 919  
Wibking B. D. et al., 2017, preprint ([arXiv:1709.07099](https://arxiv.org/abs/1709.07099))  
Wright A. H. et al., 2017, *MNRAS*, 470, 283  
Yang X., Mo H. J., van den Bosch F. C., 2008, *ApJ*, 676, 248  
Zehavi I. et al., 2011, *ApJ*, 736, 59

## APPENDIX A: ANALYTICAL COVARIANCE MATRIX

Here, we present the expressions for the covariance of the auto-correlation and cross-correlation function of two observables, in our case specifically for the galaxy–galaxy, galaxy–mass auto-correlation functions and the cross-correlation function between the two. The expressions are an extension to the Gaussian part of the covariance as presented in Singh et al. (2017) and include the non-Gaussian terms and the super-sample covariance terms that are by themselves an extension (Krause & Eifler 2017) to the non-Gaussian terms as previously described for cosmic shear only by Takada & Hu (2013) and Li et al. (2014). We follow Singh et al. (2017) and Marian et al. (2015) for the Gaussian terms, but excluding the additional contributions that arise due to not subtracting the signal obtained using random positions on the sky, as in our analysis this is performed during the signal extraction.

In general, the covariance matrix between two observables can be written as

$$\text{Cov}(X, Y) = \text{Cov}^G(X, Y) + \text{Cov}^{\text{NG}}(X, Y) + \text{Cov}^{\text{SSC}}(X, Y), \quad (\text{A1})$$

where  $X$  and  $Y$  are either  $w_p(r_p)$  or  $\Delta \Sigma_{\text{gm}}(r_p)$ , and the G stands for the Gaussian term, NG for the non-Gaussian term, and SSC stands for the contributions from the super-sample covariance. Furthermore, following Singh et al. (2017) starting with equation (A18) and Marian et al. (2015) using the derivations in their appendix, the Gaussian terms for each auto-correlation or cross-correlation can be written as (where indices  $i, j$  stand for individual projected radial bins and indices  $n, m$  stand for individual galaxy sample bins):

$$\text{Cov}^G[w_p^n, w_p^m](r_{p,i}, r_{p,j}) = 2 \frac{\mathcal{A}_{W,2}(r_{p,i}, r_{p,j})}{\mathcal{A}_{W,1}(r_{p,i}) \mathcal{A}_{W,1}(r_{p,j})} \int \frac{dk}{2\pi} J_0(kr_{p,i}) J_0(kr_{p,j}) \left( P_{\text{gg}}^n(k) + \delta_{nm} \frac{1}{\bar{n}_g^n} \right) \left( P_{\text{gg}}^m(k) + \delta_{nm} \frac{1}{\bar{n}_g^m} \right), \quad (\text{A2})$$

$$\begin{aligned} \text{Cov}^G[\Delta \Sigma_{\text{gm}}^n, \Delta \Sigma_{\text{gm}}^m](r_{p,i}, r_{p,j}) &= \bar{\rho}_m^2 \frac{\mathcal{A}_{W,2}(r_{p,i}, r_{p,j})}{\mathcal{A}_{W,1}(r_{p,i}) \mathcal{A}_{W,1}(r_{p,j})} \int \frac{dk}{2\pi} J_2(kr_{p,i}) J_2(kr_{p,j}) \left( P_{\text{gg}}^n(k) + \delta_{nm} \frac{1}{\bar{n}_g^n} \right) \left( P_{\text{mm}}^m(k) + \delta_{nm} \frac{1}{\bar{n}_\gamma^m} \right) \\ &\quad + \bar{\rho}_m^2 \frac{\mathcal{A}_{W,2}(r_{p,i}, r_{p,j})}{\mathcal{A}_{W,1}(r_{p,i}) \mathcal{A}_{W,1}(r_{p,j})} \int \frac{dk}{2\pi} J_2(kr_{p,i}) J_2(kr_{p,j}) P_{\text{gm}}^n(k) P_{\text{gm}}^m(k), \end{aligned} \quad (\text{A3})$$

$$\text{Cov}^G[w_p^n, \Delta \Sigma_{\text{gm}}^m](r_{p,i}, r_{p,j}) = \bar{\rho}_m \frac{\mathcal{A}_{W,2}(r_{p,i}, r_{p,j})}{\mathcal{A}_{W,1}(r_{p,i}) \mathcal{A}_{W,1}(r_{p,j})} \int \frac{dk}{2\pi} J_0(kr_{p,i}) J_2(kr_{p,j}) \left[ \left( P_{\text{gg}}^n(k) + \delta_{nm} \frac{1}{\bar{n}_g^n} \right) P_{\text{gm}}^n + \left( P_{\text{gg}}^m(k) + \delta_{nm} \frac{1}{\bar{n}_g^m} \right) P_{\text{gm}}^m \right], \quad (\text{A4})$$

where

$$\mathcal{A}_{W,1}(r_p) = \int \frac{dk}{2\pi} J_0(kr_p) \tilde{W}^2(k) \quad (\text{A5})$$

and

$$\mathcal{A}_{W,2}(r_{p,i}, r_{p,j}) = \int \frac{dk}{2\pi} J_0(kr_{p,i}) J_0(kr_{p,j}) \tilde{W}^2(k) \quad (\text{A6})$$

are the pre-factors arising from the survey geometry, with  $J_n$  being Bessel function of the  $n$ -th kind,  $\tilde{W}(k)$  is the window function defined in equation (A8),  $\delta_{nm}$  is the Kronecker delta symbol,  $\bar{n}_g$  the number density of galaxies in the bin  $n$ ,  $P_{\text{xy}}(k)$  are individual power spectra as defined in Section 3.1, and  $\bar{n}_\gamma$  is the shape noise given by (see also Marian et al. 2015; equation C2):

$$\frac{1}{\bar{n}_\gamma} = \frac{\sigma_\gamma^2}{\bar{n}_s} \frac{\Sigma_{\text{cr,com}}^2}{\bar{\rho}_m^2} \frac{D^2(z_1)}{2\Pi_{\text{max}}}, \quad (\text{A7})$$

where  $\sigma_\gamma$  is the shape variance of the sources used in the analysis,  $\bar{n}_s$  is the source density given by Hildebrandt et al. (2017) in gal arcmin<sup>-2</sup> (converted to radians), the  $D(z_1)$  is the angular diameter distance at  $z_1$  and  $\Pi_{\text{max}}$  is the projection length used throughout this work. The value  $\bar{n}_g$  and  $\sigma_\gamma$  are measured from the lens and source galaxies used in our three samples, respectively. We assume a circular survey geometry with a window function:

$$\tilde{W}(k) = 2\pi R^2 \frac{J_1(kR)}{kR}, \quad (\text{A8})$$

where  $R$  is the radius of the circular window with area covering 180 deg<sup>2</sup>. The  $\Sigma_{\text{cr,com}}$  is calculated using the same prescription as defined in equation (54). To project our covariance matrices, we use the Limber approximation as demonstrated by Marian et al. (2015), using the slightly more accurate survey area normalization by Singh et al. (2017).

The super-sample covariance terms are given by the following expressions:

$$\text{Cov}^{\text{SSC}}[w_p^n, w_p^m](r_{p,i}, r_{p,j}) = \frac{\mathcal{A}_{W,2}(r_{p,i}, r_{p,j})}{\mathcal{A}_{W,1}(r_{p,i}) \mathcal{A}_{W,1}(r_{p,j})} \int \frac{dk}{2\pi} J_0(kr_{p,i}) J_0(kr_{p,j}) \frac{\partial P_{\text{gg}}^n(k)}{\partial \delta_b} \frac{\partial P_{\text{gg}}^m(k)}{\partial \delta_b} \sigma_{s,nm}^2, \quad (\text{A9})$$

$$\text{Cov}^{\text{SSC}}[\Delta \Sigma_{\text{gm}}^n, \Delta \Sigma_{\text{gm}}^m](r_{p,i}, r_{p,j}) = \bar{\rho}_m^2 \frac{\mathcal{A}_{W,2}(r_{p,i}, r_{p,j})}{\mathcal{A}_{W,1}(r_{p,i}) \mathcal{A}_{W,1}(r_{p,j})} \int \frac{dk}{2\pi} J_2(kr_{p,i}) J_2(kr_{p,j}) \frac{\partial P_{\text{gm}}^n(k)}{\partial \delta_b} \frac{\partial P_{\text{gm}}^m(k)}{\partial \delta_b} \sigma_{s,nm}^2, \quad (\text{A10})$$

$$\text{Cov}^{\text{SSC}}[w_p^n, \Delta \Sigma_{\text{gm}}^m](r_{p,i}, r_{p,j}) = \bar{\rho}_m \frac{\mathcal{A}_{W,2}(r_{p,i}, r_{p,j})}{\mathcal{A}_{W,1}(r_{p,i}) \mathcal{A}_{W,1}(r_{p,j})} \int \frac{dk}{2\pi} J_0(kr_{p,i}) J_2(kr_{p,j}) \frac{\partial P_{\text{gg}}^n(k)}{\partial \delta_b} \frac{\partial P_{\text{gm}}^m(k)}{\partial \delta_b} \sigma_{s,nm}^2, \quad (\text{A11})$$

where the responses  $\frac{\partial P_{xy}(k)}{\partial \delta_b}$  are given by the following equations [following galaxy–matter and galaxy–galaxy extensions to the matter–matter only responses in Takada & Hu (2013) and Li et al. (2014) as derived by Krause & Eifler (2017)]:

$$\frac{\partial P_{gg}(k)}{\partial \delta_b} = \left( \frac{68}{21} - \frac{1}{3} \frac{d \ln k^3 P_{lin}(k)}{d \ln k} \right) I_g^1(k) I_g^1(k) P_{lin}(k) + I_{gg}^1(k, k) - 2b_{g,1} P_{gg}(k), \quad (A12)$$

$$\frac{\partial P_{gm}(k)}{\partial \delta_b} = \left( \frac{68}{21} - \frac{1}{3} \frac{d \ln k^3 P_{lin}(k)}{d \ln k} \right) I_g^1(k) I_m^1(k) P_{lin}(k) + I_{gm}^1(k, k) - b_{g,1} P_{gm}(k), \quad (A13)$$

$$\frac{\partial P_{mm}(k)}{\partial \delta_b} = \left( \frac{68}{21} - \frac{1}{3} \frac{d \ln k^3 P_{lin}(k)}{d \ln k} \right) I_m^1(k) I_m^1(k) P_{lin}(k) + I_{mm}^1(k, k), \quad (A14)$$

where we have introduced the halo model integrals  $I_x^\alpha(k)$  and  $I_{xy}^\alpha(k, k')$  as:

$$I_x^\alpha(k) = \int \mathcal{H}_x(k, M) b_{i,\alpha} n(M) dM, \quad (A15)$$

$$I_{xy}^\alpha(k, k') = \int \mathcal{H}_x(k, M) \mathcal{H}_y(k', M) b_{i,\alpha} n(M) dM. \quad (A16)$$

In the case when ‘x’ equals ‘g’,  $\mathcal{H}_g(k, M)$  is the sum of the  $\mathcal{H}_c(k, M)$  and  $\mathcal{H}_s(k, M)$  functions. Bias functions  $b_{i,\alpha}$  are either 0 if  $\alpha = 0$ , 1 if  $i = m$  and  $\alpha = 1$ , and  $\int dM \mathcal{H}_g(k, M) b_h(M) n(M) / \bar{n}_g$  if  $i = g$  and  $\alpha = 1$ . Note that these subscripts are not related to the ones used in Section 3.1. In the equations (A9), (A10), and (A11), we have also used the survey variance defined as:

$$\sigma_s^2 = \frac{1}{\mathcal{A}_{W,1}(R)} \int \frac{dk k^2}{2\pi^2} \tilde{W}^2(k) P_{lin}(k). \quad (A17)$$

The connected, non-Gaussian terms of the covariance matrix can be written as (again following the extension to the matter–matter only derivation by Krause & Eifler 2017) (see also Cooray & Sheth 2002; Takada & Jain 2009):

$$\text{Cov}^{\text{NG}}[w_p^n, w_p^m](r_{p,i}, r_{p,j}) = \frac{\mathcal{A}_{W,2}(r_{p,i}, r_{p,j})}{\mathcal{A}_{W,1}(r_{p,i}) \mathcal{A}_{W,1}(r_{p,j})} \int \frac{dk_i k_i}{\sqrt{2\pi}} \int \frac{dk_j k_j}{\sqrt{2\pi}} J_0(k_i r_{p,i}) J_0(k_j r_{p,j}) \cdot T_{gggg}^{nm}(k_i, -k_i, k_j, -k_j), \quad (A18)$$

$$\text{Cov}^{\text{NG}}[\Delta \Sigma_{gm}^n, \Delta \Sigma_{gm}^m](r_{p,i}, r_{p,j}) = \bar{\rho}_m^2 \frac{\mathcal{A}_{W,2}(r_{p,i}, r_{p,j})}{\mathcal{A}_{W,1}(r_{p,i}) \mathcal{A}_{W,1}(r_{p,j})} \int \frac{dk_i k_i}{\sqrt{2\pi}} \int \frac{dk_j k_j}{\sqrt{2\pi}} J_2(k_i r_{p,i}) J_2(k_j r_{p,j}) \cdot T_{gmgm}^{nm}(k_i, -k_i, k_j, -k_j), \quad (A19)$$

$$\text{Cov}^{\text{NG}}[w_p^n, \Delta \Sigma_p^m](r_{p,i}, r_{p,j}) = \bar{\rho}_m \frac{\mathcal{A}_{W,2}(r_{p,i}, r_{p,j})}{\mathcal{A}_{W,1}(r_{p,i}) \mathcal{A}_{W,1}(r_{p,j})} \int \frac{dk_i k_i}{\sqrt{2\pi}} \int \frac{dk_j k_j}{\sqrt{2\pi}} J_0(k_i r_{p,i}) J_2(k_j r_{p,j}) \cdot T_{gggm}^{nm}(k_i, -k_i, k_j, -k_j), \quad (A20)$$

where the individual  $T_{xyzw}$  terms are given by the combination of 1-halo matter trispectrum and (2 + 3 + 4)-halo matter trispectrum as:

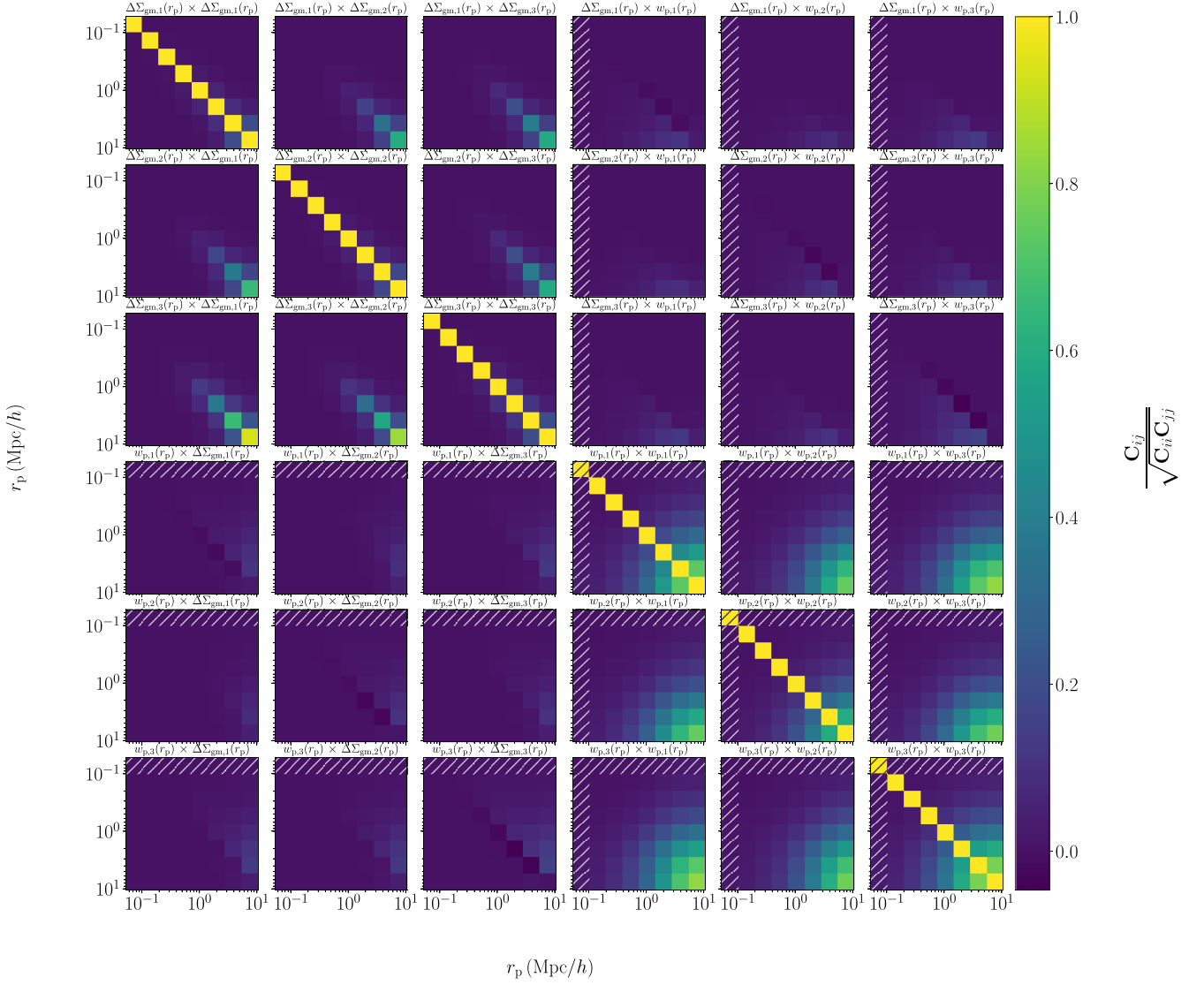
$$T_{xyzw}^{nm}(k_i, -k_i, k_j, -k_j) = (b_x b_y b_z b_w)^{nm} T_m^{2h+3h+4h}(k_i, -k_i, k_j, -k_j) + T_{xyzw}^{1h,nm}(k_i, k_i, k_j, k_j), \quad (A21)$$

where  $b_x$  is the same bias used in equations (A15) and (A16). The 1-halo matter trispectrum is given by the following integral over halo model building blocks:

$$T_{xyzw}^{1h,nm}(k_i, k_i, k_j, k_j) = \int dM n(M) \mathcal{H}_x^n(k_i, M) \mathcal{H}_y^n(k_i, M) \mathcal{H}_z^m(k_j, M) \mathcal{H}_w^m(k_j, M), \quad (A22)$$

and the (2 + 3 + 4)-halo matter trispectrum  $T_m^{2h+3h+4h}$  is calculated according to the estimate presented in Pielorz et al. (2010). Additionally, the resulting covariance matrix is bin averaged according to the same radial binning scheme as used with our data. The resulting covariance matrix obtained using the analytical prescription presented here can be seen in Fig. A1 (shown as a correlation matrix), which shows the auto-correlation between the three bins for the lensing signal, clustering signal, and the cross-correlation between the two observables in the off-diagonal  $3 \times 3$  blocks. Individual combinations between all the bins are marked above the corresponding block matrices.

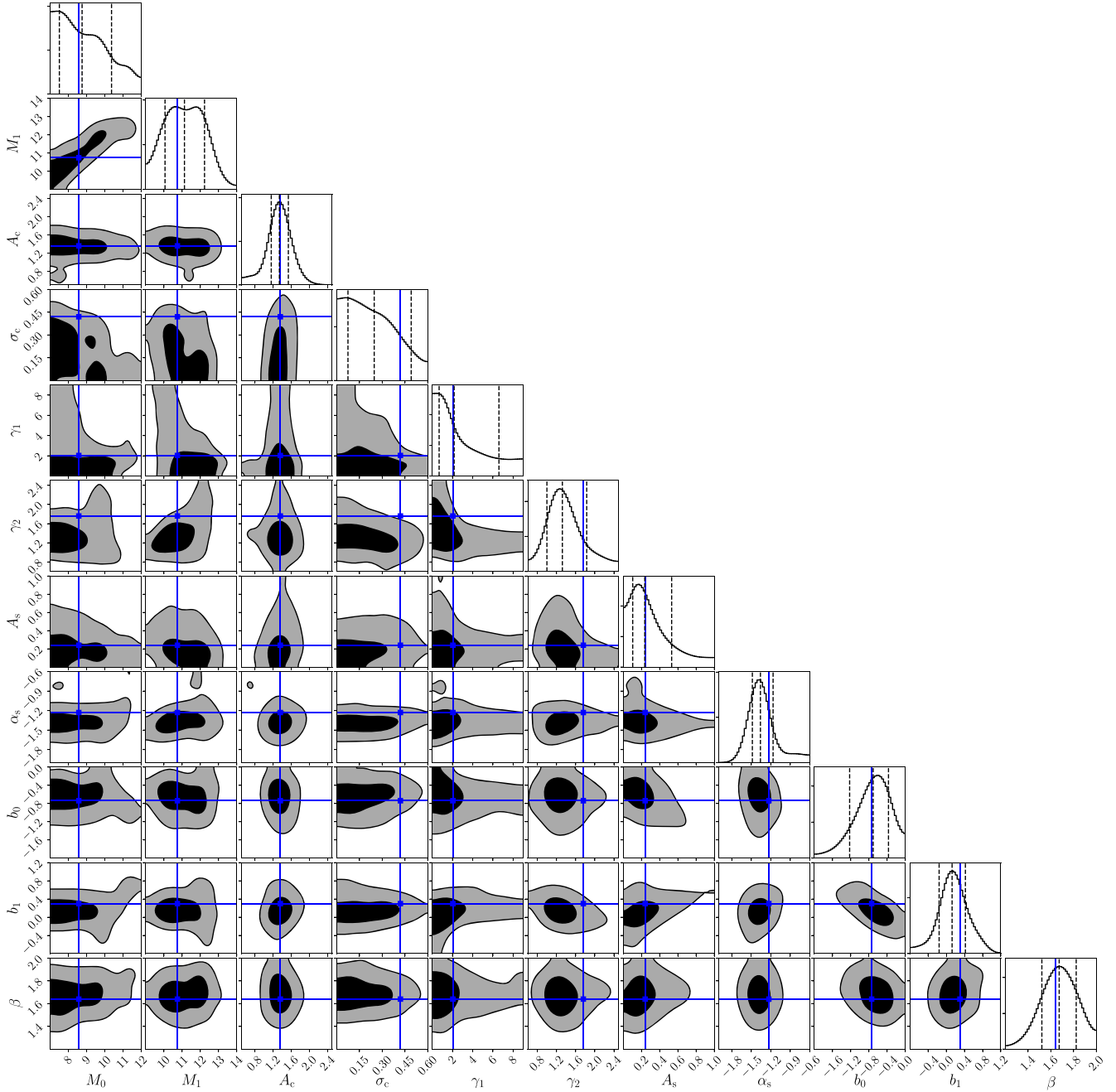




**Figure A1.** The full analytical correlation matrix for the lensing and clustering signals and their cross terms. Individual combinations between all the bins are marked above the corresponding block matrices, with indices 1, 2, and 3 corresponding to the stellar mass bins as defined in Table 1. We do not use the covariance estimates in the hatched areas in our fit, as the clustering measurements are affected by the blending on these scales.

**APPENDIX B: FULL POSTERIOR DISTRIBUTIONS**

In Fig. B1 we show the full posterior probability distribution for all fitted parameters in our MCMC fit as discussed in Section 4.



**Figure B1.** The full posterior distributions of the halo model parameters (where the priors are listed in Table 2). The contours indicate  $1\sigma$  and  $2\sigma$  confidence regions.

**APPENDIX C: RELATION BETWEEN THE GALAXY–GALAXY LENSING SIGNAL AND THE GALAXY–MATTER CROSS-CORRELATION FUNCTION**

In this appendix, we provide a step-by-step derivation of the relation between the galaxy–galaxy lensing signal and the galaxy–matter cross-correlation function. As a side-product, we motivate the two different definitions of the critical surface mass density that are used in this field. Finally, we compare our results with those in some recent papers, pointing out differences and discussing their implications. Since the results of this appendix apply to several papers, we choose to use a slightly more explicit notation here in comparison to the rest of this paper.

## C1 Derivation

The equivalent weak lensing convergence  $\kappa$  for a 3D mass distribution characterized by the fractional density contrast  $\delta$ , for sources at comoving distance  $\chi_s$ , is given by (e.g. Bartelmann & Schneider 2001; Schneider 2006)

$$\kappa(\boldsymbol{\theta}) = \frac{3H_0^2\Omega_m}{2c^2} \int_0^{\chi_s} d\chi \frac{\chi(\chi_s - \chi)}{\chi_s a(\chi)} \delta(\chi\boldsymbol{\theta}, \chi) = \bar{\rho}_m \frac{4\pi G}{c^2} \int_0^{\chi_s} d\chi \frac{\chi(\chi_s - \chi)}{\chi_s a(\chi)} \delta(\chi\boldsymbol{\theta}, \chi), \quad (C1)$$

where we assumed for notational simplicity a spatially flat cosmological model. Here,  $\bar{\rho}_m$  is the current mean matter density in the Universe, and we used the relation between mass density and density parameter in the second step, i.e.  $\bar{\rho}_m = 3H_0^2\Omega_m/(8\pi G)$ . The relation (C1) is valid in the framework of the Born approximation and by neglecting lens–lens coupling (see, e.g. Hilbert et al. 2009; Krause & Hirata 2010 for the impact of these effects).

Let  $\delta_g$  be the 3D fractional density contrast of galaxies of a given type. Their fractional density contrast on the sky,  $\kappa_g(\boldsymbol{\theta}) = [n(\boldsymbol{\theta}) - \bar{n}]/\bar{n}$ , with  $\bar{n}$  being the mean number density, is related to  $\delta_g$  by

$$\kappa_g(\boldsymbol{\theta}) = \int d\chi p_f(\chi) \delta_g(\chi\boldsymbol{\theta}, \chi), \quad (C2)$$

where  $p_f(\chi)$  is the probability distribution of the selected ‘foreground’ galaxy population in comoving distance, equivalent to a redshift probability distribution. For the following we will assume that this distribution is a very narrow one around redshift  $z_1$ , and thus approximate  $p_f(\chi) = \delta_D(\chi - \chi_1)$ . We assume throughout that  $\chi_1 < \chi_s$ . The correlator between  $\kappa$  and  $\kappa_g$  then becomes

$$\begin{aligned} \langle \kappa_g(\boldsymbol{\theta}) \kappa(\boldsymbol{\theta} + \boldsymbol{\vartheta}) \rangle &= \bar{\rho}_m \frac{4\pi G}{c^2} \int_0^{\chi_s} d\chi \frac{\chi(\chi_s - \chi)}{\chi_s a(\chi)} \int d\chi' \delta_D(\chi' - \chi_1) \langle \delta_g(\chi'\boldsymbol{\theta}, \chi') \delta[\chi(\boldsymbol{\theta} + \boldsymbol{\vartheta}), \chi] \rangle \\ &= \bar{\rho}_m \frac{4\pi G}{c^2} \int_0^{\chi_s} d\chi \frac{\chi(\chi_s - \chi)}{\chi_s a(\chi)} \langle \delta_g(\chi_1\boldsymbol{\theta}, \chi_1) \delta[\chi(\boldsymbol{\theta} + \boldsymbol{\vartheta}), \chi] \rangle. \end{aligned} \quad (C3)$$

Since the correlator is significantly non-zero only over a small interval in  $\chi$  around  $\chi_1$ , the prefactor in the integrand can be considered to be constant over this interval and taken out of the integral. This yields

$$\begin{aligned} \langle \kappa_g(\boldsymbol{\theta}) \kappa(\boldsymbol{\theta} + \boldsymbol{\vartheta}) \rangle &= \frac{4\pi G}{c^2} \frac{\chi_1(\chi_s - \chi_1)}{\chi_s a(\chi_1)} \bar{\rho}_m \int_0^{\chi_s} d\chi \xi_{gm} \left( \sqrt{\chi_1^2 |\boldsymbol{\vartheta}|^2 + (\chi - \chi_1)^2} \right) \\ &= \Sigma_{cr,com}^{-1} \Sigma_{com}(\chi_1 |\boldsymbol{\theta}|), \end{aligned} \quad (C4)$$

where the galaxy–matter cross-correlation function  $\xi_{gm}$  (at fixed redshift  $z_1$ ) is defined through

$$\langle \delta(\mathbf{x}) \delta_g(\mathbf{x} + \mathbf{y}) \rangle = \xi_{gm}(|\mathbf{y}|), \quad (C5)$$

in which  $\mathbf{x}$  and  $\mathbf{y}$  are comoving spatial vectors, and the sole dependence on  $|\mathbf{y}|$  is due to the assumed homogeneity and isotropy of the density fields in the Universe. Furthermore, we have defined the comoving critical surface mass density  $\Sigma_{cr,com}$  through

$$\Sigma_{cr,com}^{-1} = \frac{4\pi G}{c^2} \frac{\chi_1(\chi_s - \chi_1)}{\chi_s a(\chi_1)} H(\chi_s - \chi_1), \quad (C6)$$

with  $H(x)$  being the Heaviside unit step function,<sup>4</sup> and the comoving surface mass density as

$$\Sigma_{com}(R_{com}) = \bar{\rho}_m \int_0^{\chi_s} d\chi \xi_{gm} \left( \sqrt{R_{com}^2 + (\chi - \chi_1)^2} \right), \quad (C7)$$

as a function of the comoving projected separation  $R_{com}$ . In this paper,  $\Sigma_{cr,com}$  is termed  $\Sigma_{crit}$  – see equation (50), and  $R_{com}$  and  $\Sigma_{com}$  are called  $r_p$  and  $\Sigma_{gm}$  – see equation (43).

The interpretation of equation (C4) is then that  $\bar{\rho}_m \xi_{gm}$  is the average comoving overdensity of matter around galaxies, caused by the correlation between them, and that the integral over comoving distance then yields the comoving surface mass density of this excess matter. The corresponding convergence is then obtained by scaling with the comoving critical surface mass density  $\Sigma_{cr,com}$ .

There is another form in which equation (C4) can be written by rearranging factors of  $a(\chi_1)$ , namely

$$\begin{aligned} \langle \kappa_g(\boldsymbol{\theta}) \kappa(\boldsymbol{\theta} + \boldsymbol{\vartheta}) \rangle &= \frac{4\pi G}{c^2} \frac{\chi_1(\chi_s - \chi_1) a(\chi_1)}{\chi_s} \bar{\rho}_m a^{-2}(\chi_1) \int_0^{\chi_s} d\chi \xi_{gm} \left( \sqrt{\chi_1^2 |\boldsymbol{\vartheta}|^2 + (\chi - \chi_1)^2} \right) \\ &=: \Sigma_{cr}^{-1} \Sigma(\chi_1 |\boldsymbol{\theta}|), \end{aligned} \quad (C8)$$

where we defined the (proper) critical surface mass density  $\Sigma_{cr}$  through

$$\Sigma_{cr}^{-1} = \frac{4\pi G}{c^2} \frac{\chi_1(\chi_s - \chi_1) a(\chi_1)}{\chi_s} H(\chi_s - \chi_d) = \frac{4\pi G}{c^2} \frac{D_1 D_{ls}}{D_s}, \quad (C9)$$

<sup>4</sup>The corresponding expression for a general curvature parameter reads

$$\Sigma_{cr,com}^{-1} = \frac{4\pi G}{c^2} \frac{f_K(\chi_1) f_K(\chi_s - \chi_1)}{f_K(\chi_s) a(\chi_1)} H(\chi_s - \chi_1),$$

where  $f_K(\chi)$  is the comoving angular-diameter distance to a comoving distance  $\chi$ , and either the identity for spatially flat models, or a sin or sinh function for closed or open models, respectively.

and in the last step we introduced the angular-diameter distances  $D_l = D(0, z_l)$ ,  $D_s = D(0, z_s)$ , and  $D_{ds} = D(z_l, z_s)$ , with  $D(z_1, z_2) = a(z_2) [\chi(z_2) - \chi(z_1)] H(z_2 - z_1)$  being the angular-diameter distance of a source at redshift  $z_2$  as seen from an observer at redshift  $z_1$ .<sup>5</sup> Furthermore,

$$\Sigma(R_{\text{com}}) = \rho_m(\chi_l) \int_0^{\chi_s} d\chi a(\chi) \xi_{\text{gm}} \left( \sqrt{R_{\text{com}}^2 + (\chi - \chi_l)^2} \right). \quad (\text{C10})$$

We note that, due to the assumed localized nature of the correlation function, we could write  $a(\chi)$  into the integrand in equation (C10). The interpretation of equation (C8) is now that the (proper) overdensity around galaxies caused by the galaxy–matter cross-correlation,  $\rho_m(\chi_l) \xi_{\text{gm}} = \bar{\rho}_m(1 + z_l)^3 \xi_{\text{gm}}$ , is integrated along the l.o.s. in proper coordinates,  $dr_{\text{prop}} = a d\chi$ , and the resulting (proper) surface mass density is scaled by the critical surface mass density  $\Sigma_{\text{cr}}$ . We note that the argument of the (proper) surface mass density  $\Sigma$  is a comoving transverse separation, since the correlation function is a function of comoving separation.

The relation between the two different equations (C6) and (C9) of the critical surface mass density is

$$\Sigma_{\text{cr,com}} = a^2(\chi_l) \Sigma_{\text{cr}}, \quad (\text{C11})$$

so that the comoving critical surface density is smaller by a factor  $a^2(\chi_l)$ . This makes sense: for a given lens, the comoving surface mass density (mass per unit comoving area) is smaller than the proper surface mass density

$$\Sigma_{\text{com}}(R_{\text{com}}) = a^2(\chi_l) \Sigma(R_{\text{com}}), \quad (\text{C12})$$

since the comoving area is larger than the proper one by a factor  $a^{-2}$ . Correspondingly, since the convergence, or the correlation function in equation (C4), is independent of whether proper or comoving measures are used, the comoving critical surface mass density is smaller by the same factor.

If  $N$  lensing galaxies at redshift  $z_l$  are located at positions  $\theta_i$  within a solid angle  $\omega$ , the corresponding fractional number density contrast reads

$$\kappa_g(\theta) = \frac{1}{\bar{n}} \sum_{i=1}^N \delta_D(\theta - \theta_i) - 1, \quad (\text{C13})$$

where for large  $N$  and  $\omega$ ,  $\bar{n} = N/\omega$ . To evaluate the correlator of equation (C3) in this case, we replace the ensemble average with an angular average, as is necessarily done in any practical estimation

$$\langle \kappa_g(\theta) \kappa(\theta + \vartheta) \rangle \approx \frac{1}{\omega} \int_{\omega} d^2\theta \kappa_g(\theta) \kappa(\theta + \vartheta) = \frac{1}{\omega} \int_{\omega} d^2\theta \left[ \frac{1}{\bar{n}} \sum_{i=1}^N \delta_D(\theta - \theta_i) \right] \kappa(\theta + \vartheta) = \frac{1}{N} \sum_{i=1}^N \kappa(\theta_i + \vartheta), \quad (\text{C14})$$

valid for separations  $\vartheta$  which are much smaller than the linear angular extent  $\sqrt{\omega}$  of the region (to neglect boundary effects), and we employed the fact that the ensemble average – and in the same approximation as above, the angular average – of  $\kappa(\vartheta)$  vanishes. We thus see that the correlator  $\kappa_g \kappa$  can be obtained from the average convergence around the foreground galaxies, a quantity probed by the shear. Thus, we find the relations

$$\gamma_l(\theta) = \Sigma_{\text{cr}}^{-1} \Delta \Sigma(\chi_l \theta) = \Sigma_{\text{cr,com}}^{-1} \Delta \Sigma_{\text{com}}(\chi_l \theta), \quad (\text{C15})$$

where

$$\Delta \Sigma(R_{\text{com}}) = \frac{2}{R_{\text{com}}^2} \int_0^{R_{\text{com}}} dR R \Sigma(R) - \Sigma(R_{\text{com}}), \quad (\text{C16})$$

and the analogous definition for  $\Delta \Sigma_{\text{com}}$ .

A further subtlety and potential source of confusion is that frequently, the surface mass density  $\Sigma$  is considered a function of *proper* transverse separation  $R = a(\chi_l) R_{\text{com}}$ . For the purpose of this appendix, we call this function  $\Sigma_p$ , which is related to  $\Sigma$  by

$$\Sigma_p(R) = \Sigma[R/a(\chi_l)], \quad \text{or} \quad \Sigma(R_{\text{com}}) = \Sigma_p[a(\chi_l) R_{\text{com}}], \quad (\text{C17})$$

yielding

$$\gamma_l(\theta) = \Sigma_{\text{cr}}^{-1} \Delta \Sigma(\chi_l \theta) = \Sigma_{\text{cr}}^{-1} \Delta \Sigma_p(D_l \theta). \quad (\text{C18})$$

We argue that the definition used should depend on the science case. For example, when considering the mean density profile of galaxies, it is more reasonable to use proper transverse separations – as that density profile is expected to be approximately stationary in proper coordinates. For larger-scale correlations between galaxies and matter, however, the use of comoving transverse separations is more meaningful, since the shape of the cross-correlation function on large scales is expected to be approximately preserved.

## C2 Relation to previous work

In the literature on galaxy–galaxy lensing, one finds relations that differ from the ones derived above; we shall comment on some of these differences here.

<sup>5</sup>For a model with free curvature,  $D(z_1, z_2) = a(z_2) \{f_K [\chi(z_2) - \chi(z_1)]\} H(z_2 - z_1)$ .



The first aspect is that in several papers (e.g. Mandelbaum et al. 2010; Viola et al. 2015; de la Torre et al. 2017), the integrand in equation (C4) is replaced by  $1 + \xi_{\text{gm}}$ , implying that the corresponding  $\Sigma_{\text{com}}$  contains the line-of-sight integrated mean density of the Universe, in addition to the correlated density. This constant term is, however, not justified by the derivation in Appendix C1. While such a constant drops out in the definition of  $\Delta \Sigma_{\text{com}}$ , and thus does not impact on quantitative results, it nevertheless causes a principal flaw: its inclusion would imply that the convergence  $\kappa = \Sigma_{\text{cr,com}}^{-1} \Sigma_{\text{com}}$  for all lines-of-sight to redshifts  $z_s \sim 1$  would be several tenths, causing a large difference between shear and reduced shear [ $g = \gamma/(1 + \kappa)$ ], which is the observable in weak lensing. That is, it would strongly modify the relation between  $\gamma$  and the observable image ellipticities, yielding significant biases in all weak lensing studies. Indeed, the mean density of the Universe is already taken into account by the Robertson–Walker metric: the fact that the angular-diameter distance is a non-monotonic function of redshift can be considered as being due to the gravitational light deflection by the mean mass density of the Universe – the convergence part in the optical tidal equation (see e.g. Seitz, Schneider & Ehlers 1994). However, this is usually not called ‘lensing’, but ‘curvature of the metric’. Lensing is usually ascribed solely to the effect caused by density inhomogeneities. But in any case: the mean cosmic density cannot be accounted for twice, once for the metric [and thus the use of (comoving) angular-diameter distances in a FRW model], and a second time for the convergence on such a background model.

A second issue in some of the recent GGL papers is mixing the use of the comoving surface mass density,  $\Sigma_{\text{com}}$  (equation C7), with the proper critical surface mass density,  $\Sigma_{\text{cr}}$  (equation C9). For example, de la Torre et al. (2017) in their equation (18) (apart from the constant term discussed above) define  $\Sigma_{\text{com}}$  as in equation (C9), but use in their equations (9) and (10) the proper critical surface mass density  $\Sigma_{\text{cr}}$  to relate the tangential shear to  $\Sigma_{\text{com}}$ . Hence, this relation would cause an offset by a factor  $(1 + z_1)^2$  from the correct result. However, the inconsistency appears only in the text of the paper and not in the code or calculations performed (de la Torre, private communication).

The same issue occurs in the write-up in several earlier publications of our KiDS team. For example, the equations (2, 5, and 6) in Viola et al. (2015) show this inconsistency (where we also point out a typo in the integration limits of equation 2), as well as equations (1, 2, 3) in van Uitert et al. (2016) and, as mentioned already in the main text, equations (1, 2, 3) in Dvornik et al. (2017). We have checked the codes that were used to derive the quantitative results in these papers to see whether they employ the same inconsistent use of quantities. We found that there is an inconsistency present only in writing, namely in equation (2) of Viola et al. (2015) and not in the code that was used to produce the results. As pointed out above the correlation function is given in comoving coordinates, while the extraction of the galaxy–galaxy lensing signal is calculated in proper coordinates. Thus, the equation (2) of Viola et al. (2015) should correctly read as (using their notation):

$$\Sigma(R) = 2\bar{\rho}_m (1 + \langle z_1 \rangle)^2 \int_0^{\pi_s} \xi_{\text{gm}}(\sqrt{R^2 + \Pi^2}) d\Pi. \quad (\text{C19})$$

Put differently, while the data were indeed extracted in proper coordinates, the outputs of the halo model were not reported to be scaled to the coordinates used by the data (Cacciato private communication). Secondly, the equations (6) and (9) in Dvornik et al. (2017) should have the same form as equation (50) and (54) in this paper, again having a mistake only in writing. The exact same correction should also be applied to the equation (3) of Velliscig et al. (2017).

The erroneous formulation also occurred in a previous paper from van Uitert et al. (2016). Their equation (2), should read correctly as equation (50), which is (again, using their notation):

$$\Sigma_{\text{crit}} = \frac{c^2}{4\pi G(1+z)^2} \frac{D_s}{D_L D_{\text{LS}}}. \quad (\text{C20})$$

Further KiDS analyses from Amon et al. (2018a) and Amon et al. (2018b) measure large-scale galaxy–mass correlations using the comoving critical surface mass density. The definitions, presented in these two papers are consistent with the data and the equations derived in Appendix C1.

On the other hand, the KiDS analyses from Sifón et al. (2015), Brouwer et al. (2016), and Brouwer et al. (2017) use a separate NFW stacking method that does not rely on the halo model, and all use proper coordinates that are consistent with the data.

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.